



(Cahiers Mathématiques de l'Université de Sherbrooke)

Revue : CaMUS (Cahiers Mathématiques de l'Université de Sherbrooke)

Volume : 6

Année : 2018

Éditeur : Université de Sherbrooke. Département de Mathématiques

URI : Repéré à : <http://camus.math.usherbrooke.ca/revue.html>

Page vide laissée intentionnellement

CAMUS

Volume 6 Cahiers mathématiques de
l'Université de Sherbrooke

M. Frappier	1
<i>Itérations d'un processus avec mémoire défini par l'application tente</i>	
T. Sánchez McMillan	13
<i>Classification des représentations indécomposables du carquois de Kronecker</i>	
K. Thouin	29
<i>Géométrie hyperbolique : le demi-plan de Poincaré</i>	
S. Hassoun, D. Langford et F. Langlois	45
<i>La structure d'algèbre amassée des grassmaniennes $Gr(k, n)$</i>	
V. Roy	69
<i>Dégénération des représentations des carquois</i>	
S. Goyette	81
<i>Analyse fine d'algorithmes de type ARC-TR au cas unidimensionnel</i>	

Itérations d'un processus avec mémoire défini par l'application tente

Mathieu Frappier

Reçu le 2016-06-27 et accepté le 2016-10-02

RÉSUMÉ Dans cet article, on étudie le processus avec mémoire défini par : $x_{n+1} = f(\alpha x_n + (1 - \alpha)x_{n-1})$ où f est l'application tente, $0 < \alpha < 1$ et $x_i \in [0,1]$. On voit que dans le cas $\alpha = \frac{1}{2}$, tous les points sont de période trois ou le seront éventuellement sauf pour deux points fixes. On termine en présentant des conjectures énoncées par P. Góra dont certaines qui ne sont pas prouvées actuellement pour $\alpha \neq \frac{1}{2}$.

1 Introduction

L'étude des systèmes dynamiques intéresse les mathématiciens depuis le dix-septième siècle. Au départ, Isaac Newton s'y intéressait pour décrire des processus qui évoluent dans le temps. Ils étaient souvent modélisés par des équations différentielles. Au fil des siècles, certains mathématiciens ont fait des découvertes importantes. Par exemple, en 1963 Edward Lorenz a modélisé des phénomènes météorologiques à l'aide d'un système d'équations différentielles. Il a découvert ce que l'on appelle l'attracteur de Lorenz. Cet attracteur chaotique a donné naissance à l'effet papillon. Ce concept n'a pas seulement été apprécié par les mathématiciens mais par la population en général.

Aujourd'hui, on peut rapidement calculer les itérations d'une fonction et on peut aussi visualiser facilement l'évolution d'un système avec les nouvelles technologies. Bien qu'on possède ces outils, l'étude des systèmes dynamiques est toujours pertinente : On peut y trouver des applications dans plusieurs domaines dont en physique, en écologie, en finance, etc.

2 Définitions

Commençons par énoncer quelques définitions élémentaires tirées des livres de Devaney [Dev92], [Dev03] et reprises dans [Rob95].

Je tiens à remercier le CRSNG pour son financement et M. Tomasz Kaczynski professeur à l'Université de Sherbrooke pour la supervision de mon stage de recherche et pour l'aide qu'il a apportée à la rédaction de cet article.

Définition 2.1. Un *système dynamique* est un ensemble de composants en interaction. Il est souvent régi par un système d'équations différentielles décrivant leur mouvement.

Dans notre cas, le système dynamique étudié est la suite d'itérations d'une fonction.

Intéressons nous à la fonction $F(x) = 2\cos(x + 3)$. Notons sa deuxième itération $F^2(x) = 2\cos(2\cos(x + 3) + 3)$ et sa troisième itération $F^3(x) = 2\cos(2\cos(2\cos(x + 3) + 3) + 3)$. En général, on note $F^n(x)$ la n -ième itération de $F(x)$. Prendre note que l'exposant ne signifie pas que l'on prend la n -ième puissance de $F(x)$ ni sa n -ième dérivée.

Définition 2.2. Soit $x_0 \in \mathbb{R}$, on définit l'*orbite* de x_0 sur F comme la suite de points $x_0, x_1 = F(x_0), x_2 = F^2(x_0), \dots, x_n = F^n(x_0), \dots$

Le point x_0 est appelé la *graine* de l'*orbite*.

Définition 2.3. Soit $x_0 \in \mathbb{R}$, x_0 est un *point fixe* de $F \Leftrightarrow F(x_0) = x_0$.

Définition 2.4. Soit $x_0 \in \mathbb{R}$, x_0 est un *point éventuellement fixe* de

$$F \Leftrightarrow \exists n \in \mathbb{N} \setminus \{0\}$$

tel que

$$F^n(x_0) = F^{n+1}(x_0).$$

Définition 2.5. Un *cycle* est une orbite telle que pour $x_0 \in \mathbb{R}$, $F^n(x_0) = x_0$ avec $n > 0$. La plus petite valeur de n est appelée la *période*.

Exemple 2.6. Soit $F(x) = -x^2 + 3x + 4$, on choisit $x_0 = -1$. Calculons l'orbite de x_0 . $x_1 = F(x_0) = 0$, $x_2 = F(x_1) = 4$, $x_3 = F(x_2) = 0$. Donc ici, l'orbite de x_0 est la suite -1, 0, 4, 0, 4, 0, ... Elle devient cyclique de période deux.

3 Étude du système

Dans cette section, on présente la solution à un problème qui a été soumis par M. Pawel Góra, professeur à l'Université Concordia, lors de la réunion de la Société mathématique du Canada en juin 2015. J'y ai travaillé lors d'un stage de recherche supervisé par M. Tomasz Kaczynski. M. Góra et ses collègues ont aussi apporté des solutions de leur côté [GBLP16]. Le problème est :

Pour une certaine valeur $0 < \alpha < 1$, on considère le processus avec mémoire défini par $x_{n+1} = f(\alpha x_n + (1 - \alpha)x_{n-1})$. On dit que x_0 est l'état initial. x_n représente l'état actuel du processus, x_{n-1} représente sa plus récente valeur passée et x_{n+1} représente sa prochaine valeur future. Plus précisément, on s'intéresse au cas où f est l'application tente.

$$f(x) = \begin{cases} \mu x & \text{si } x < \frac{1}{2} \\ \mu(1 - x) & \text{si } x \geq \frac{1}{2} \end{cases}$$

On choisit cette application car elle possède plusieurs propriétés intéressantes dans l'étude des systèmes dynamiques chaotiques. En particulier, on peut prouver qu'il existe des cycles de période n pour tout $n \in \mathbb{N} \setminus \{0\}$. On prend $\mu = 2$ pour simplifier les calculs et on remarque que le système se comporte de manière très différente selon le paramètre α . La prochaine tâche est de montrer que pour $\alpha = \frac{1}{2}$, la suite des itérations de x_n sera éventuellement de période trois sauf pour $x_n = 0$ et $x_n = \frac{2}{3}$.

Pour ce faire, on pose la fonction G et on regarde ses itérations.

$$\begin{aligned} G &: [0,1]^2 \rightarrow [0,1]^2 \\ G(x,y) &\mapsto (y, f(\alpha y + (1-\alpha)x)) \end{aligned}$$

$$(x_n, y_n) = G(x_{n-1}, y_{n-1}) \Leftrightarrow x_n = y_{n-1}$$

et

$$\begin{aligned} y_n &= f(\alpha y_{n-1} + (1-\alpha)x_{n-1}) \\ &= f(\alpha x_n + (1-\alpha)x_{n-1}) \end{aligned}$$

mais

$$x_{n+1} = y_n \Rightarrow x_{n+1} = f(\alpha x_n + (1-\alpha)x_{n-1}).$$

Donc, le problème initial est équivalent aux itérations de l'application G . Ils vont se comporter de manière semblable. L'avantage avec la fonction G est qu'il est possible de passer d'un problème de degré deux dans \mathbb{R} à un problème de degré un dans \mathbb{R}^2 .

Par la suite, posons : $\alpha = \frac{1}{2}$, $v_0 = (x_0, y_0)$ avec $v_0 \in [0,1]^2$. Notons $v_1 = G(v_0)$, $v_2 = G(v_1) = G^2(v_0)$, \dots , $v_n = (x_n, y_n) = G^n(v_0)$. On constate que si $x_i + y_i < 1$, on prend $f(x) = 2x$ dans l'application. Sinon on choisit $f(x) = 2(1-x)$.

Remarquons aussi que $(0,0)$ est un point fixe. C'est-à-dire que pour $v = (0,0)$, on a $G(v) = (0,0)$. Commençons par montrer les propositions suivantes.

Proposition 3.1. *Si $x_i + y_i \geq 1$, alors $x_{i+n} + y_{i+n} \geq 1 \forall n \in \mathbb{N}$.*

Démonstration. Si $x_i + y_i \geq 1$, on a $\frac{x_i + y_i}{2} \geq \frac{1}{2}$

$$\begin{aligned} \Rightarrow G(v_i) &= \left(y_i, f\left(\frac{1}{2}y_i + \frac{1}{2}x_i\right) \right) \\ &= (y_i, 2 - (x_i + y_i)) \end{aligned}$$

$$\begin{aligned} \Rightarrow x_{i+1} + y_{i+1} &= y_i + 2 - (x_i + y_i) \\ &= 2 - x_i \\ &\geq 1. \end{aligned}$$

□

Proposition 3.2. *Soit $v_0 \neq (0,0)$ et $n \in \mathbb{N} \setminus \{0\}$. Si $x_n + y_n < 1$ et $1 \leq i \leq n$ alors $x_i + y_i < x_{i+1} + y_{i+1}$.*

Démonstration. Soit $v_0 \neq (0, 0)$ tel que $x_i + y_i < 1 \forall i = 0, 1, \dots, n$. On a que

$$\begin{bmatrix} x_{i+1} \\ y_{i+1} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_i \\ y_i \end{bmatrix} \text{ pour } i = 0, 1, \dots, n$$

par la définition de G lorsque $x_i + y_i < 1$. Donc, on a que

$$\begin{bmatrix} x_{i+1} \\ y_{i+1} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}^{i+1} \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} \text{ pour } i = 0, 1, \dots, n.$$

On observe par récurrence que

$$\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}^{i+1} = \begin{bmatrix} F_i & F_{i+1} \\ F_{i+1} & F_{i+2} \end{bmatrix} \text{ pour } i = 0, 1, \dots, n.$$

Où F_i est le i -ième terme de la suite de Fibonacci. On rappelle que la suite de Fibonacci est définie comme étant :

$$F_{i+2} = F_{i+1} + F_i, \quad F_0 = 0, \quad F_1 = 1, \quad i \geq 0.$$

Pour $i = 1$, le résultat est trivial. On suppose que c'est vérifié pour tout i tel que $i = 0, 1, \dots, n$. On a

$$\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}^{(i+1)+1} = \begin{bmatrix} F_i & F_{i+1} \\ F_{i+1} & F_{i+2} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} F_{i+1} & F_{(i+1)+1} \\ F_{(i+1)+1} & F_{(i+1)+2} \end{bmatrix}.$$

Ainsi,

$$\begin{bmatrix} x_{i+1} \\ y_{i+1} \end{bmatrix} = \begin{bmatrix} F_i & F_{i+1} \\ F_{i+1} & F_{i+2} \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} \text{ pour } i = 0, 1, \dots, n.$$

Donc,

$$(x_{i+1} + y_{i+1}) - (x_i + y_i) = F_i x_0 + F_{i+1} y_0.$$

D'où

$$(x_{i+1} + y_{i+1}) - (x_i + y_i) > 0 \forall i = 0, 1, \dots, n \text{ si } y_0 > 0$$

et

$$(x_{i+1} + y_{i+1}) - (x_i + y_i) > 0 \forall i = 1, \dots, n \text{ si } y_0 = 0$$

car dans ce cas $x_0 > 0$. □

Remarque 3.3. Pour la proposition 3.2, on doit prendre $i \geq 1$ car dans le cas où $y_0 = 0$, on obtient

$$\begin{aligned} G(v_0) &= G((x_0, 0)) \\ &= \left(0, f\left(\frac{1}{2}x_0\right)\right) \\ &= (0, x_0) \\ &= v_1. \end{aligned}$$

Alors, $x_0 + y_0 = x_1 + y_1 = x_0$. Par contre, dès la prochaine itération on a bien $x_1 + y_1 < x_2 + y_2$.

Proposition 3.4. *Pour tout $v_0 \neq (0, 0)$, il existe $n \in \mathbb{N}$ tel que $x_n + y_n \geq 1$*

Démonstration. Le résultat est trivial par la Proposition 3.2. La suite de Fibonacci tend vers l'infini, alors on prend n aussi grand qu'on veut. \square

Remarque 3.5. Lorsqu'on atteint n tel que $x_n + y_n \geq 1$, il n'est plus possible de se servir de la suite de Fibonacci pour définir x_{n+1} .

Les Figures 1 et 2 montrent qu'après un certain nombre d'itérations le système passe au-dessus de la droite $x + y = 1$.

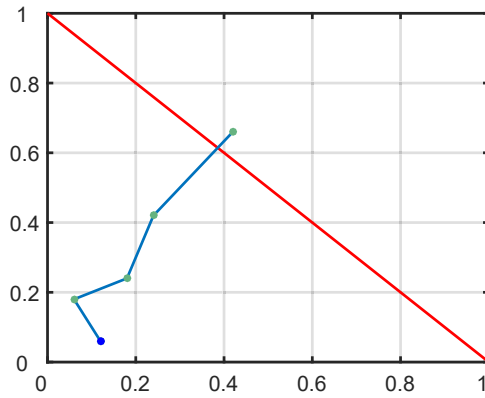


Figure 1 – Itérations du point $(0,12; 0,06)$. Après quatre itérations le système passe au-dessus de la droite $x + y = 1$.

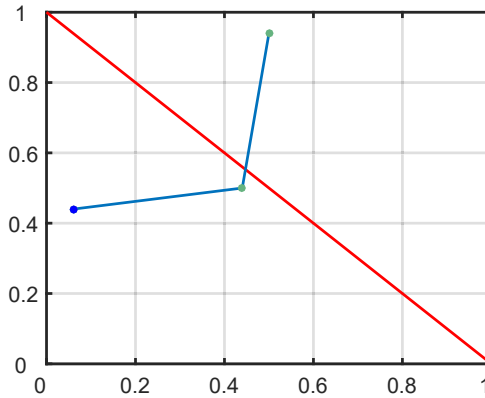


Figure 2 – Itérations du point $(0,06; 0,44)$. Ici, après deux itérations le système passe au-dessus de la droite $x + y = 1$.

Jusqu'à maintenant, on sait que tous les points sous la droite $x + y = 1$ dans $[0,1]^2$ atteindront celle-ci ou passeront au-dessus après un nombre d'itérations assez grand sauf pour le cas du point $(0, 0)$. De plus, une fois au-dessus de la droite il est impossible qu'une itération retourne sous celle-ci. Il ne reste plus qu'à itérer les points $(x_i, y_i) \in [0,1]^2$ tels que $x_i + y_i \geq 1$.

Soit $v_0 = (x_0, y_0)$ tel que $x_0 + y_0 \geq 1$.

$$v_1 = G(v_0) = (y_0, 2 - (x_0 + y_0))$$

$$\begin{aligned} v_2 &= \left(2 - (x_0 + y_0), f\left(\frac{2 - x_0}{2}\right) \right) \\ &= (2 - (x_0 + y_0), x_0) \end{aligned}$$

$$\begin{aligned} v_3 &= \left(x_0, f\left(\frac{2 - y_0}{2}\right) \right) \\ &= (x_0, y_0) \\ &= v_0. \end{aligned}$$

On vérifie que 3 est le plus petit n tel que $G^n(v_0) = v_0$.

$$\begin{aligned} v_0 = G(v_0) &\Leftrightarrow (x_0, y_0) = (y_0, 2 - (x_0 + y_0)) \\ &\Leftrightarrow \begin{cases} x_0 = y_0 \\ y_0 = 2 - (x_0 + y_0) \end{cases} \\ &\Leftrightarrow \begin{cases} x_0 = y_0 \\ x_0 = 2(1 - y_0) \end{cases} \\ &\Leftrightarrow x_0 = y_0 = \frac{2}{3} \\ &\Rightarrow v_0 = \left(\frac{2}{3}, \frac{2}{3}\right) \text{ est un point fixe.} \end{aligned}$$

$$\begin{aligned} v_0 = G^2(v_0) &\Leftrightarrow (x_0, y_0) = (2 - (x_0 + y_0), x_0) \\ &\Leftrightarrow \begin{cases} y_0 = x_0 \\ x_0 = 2 - (x_0 + y_0) \end{cases} \\ &\Leftrightarrow x_0 = y_0 = \frac{2}{3}. \end{aligned}$$

Donc les points au-dessus de la droite $x + y = 1$ sont tous de périodes trois sauf pour le point fixe $\left(\frac{2}{3}, \frac{2}{3}\right)$.

La Figure 3 montre un exemple de point qui une fois que le système passe au-dessus de la droite $x + y = 1$, il devient de période trois.

Proposition 3.6. *Les seuls points éventuellement fixes à $\left(\frac{2}{3}, \frac{2}{3}\right)$ sont $\left(\frac{2}{3}, 0\right)$ et $\left(0, \frac{2}{3}\right)$.*

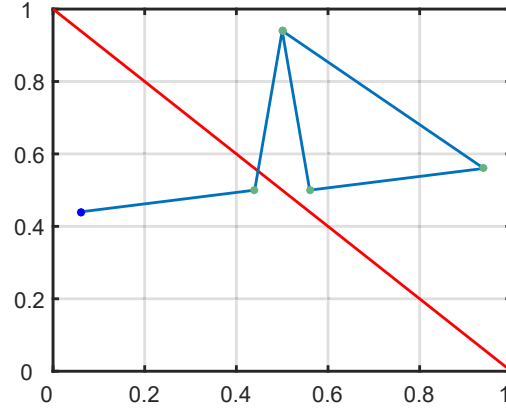


Figure 3 – Dans la Figure 2, on itère deux fois le point $x_0 = (0,06; 0,44)$ et on termine au point $(0,5; 0,94)$. Ici, on itère trois fois de plus et on termine encore au point $(0,5; 0,94)$.

Démonstration. Soit $v = (x, y)$.

Cas $x + y \geq 1$:

$$\begin{aligned}
 G(v) = (y, 2 - (x + y)) = \left(\frac{2}{3}, \frac{2}{3}\right) &\Leftrightarrow \begin{cases} y = \frac{2}{3} \\ 2 - (x + y) = \frac{2}{3} \end{cases} \\
 &\Leftrightarrow x = y = \frac{2}{3} \\
 &\Rightarrow \text{Pas de point éventuellement fixe} \\
 &\quad \text{pour } x + y \geq 1.
 \end{aligned}$$

Cas $x + y < 1$:

$$\begin{aligned}
 G(v) = (y, x + y) = \left(\frac{2}{3}, \frac{2}{3}\right) &\Leftrightarrow \begin{cases} y = \frac{2}{3} \\ x + y = \frac{2}{3} \end{cases} \\
 &\Leftrightarrow v = \left(0, \frac{2}{3}\right).
 \end{aligned}$$

Préimage de $\left(0, \frac{2}{3}\right)$:

$$\begin{aligned}
 G(v) = \left(0, \frac{2}{3}\right) &\Leftrightarrow \begin{cases} y = 0 \\ x + y = \frac{2}{3} \end{cases} \\
 &\Leftrightarrow v = \left(\frac{2}{3}, 0\right).
 \end{aligned}$$

Préimage de $\left(\frac{2}{3}, 0\right)$:

$$\begin{aligned} G(v) = \left(\frac{2}{3}, 0\right) &\Leftrightarrow \begin{cases} y = \frac{2}{3} \\ x + y = 0 \end{cases} \\ &\Leftrightarrow v = \left(-\frac{2}{3}, \frac{2}{3}\right). \end{aligned}$$

Ce qui est une contradiction car $v \in [0,1]^2$. □

4 Conclusion

Comme on a montré au début que la fonction G et le système initial défini par $x_{n+1} = f(\alpha x_n + (1 - \alpha)x_{n-1})$ sont équivalents, on en tire la même conclusion. Pour $\alpha = \frac{1}{2}$, on a deux uniques points fixes en $(0, 0)$ et $\left(\frac{2}{3}, \frac{2}{3}\right)$. On a deux points éventuellement fixes à $\left(\frac{2}{3}, \frac{2}{3}\right)$ soient $\left(0, \frac{2}{3}\right)$ et $\left(\frac{2}{3}, 0\right)$. Tous les autres points $(x, y) \in [0,1]^2$ sont de période trois ou le seront éventuellement.

D'autres conjectures pour $\alpha \neq \frac{1}{2}$ ne sont toujours pas démontrées et certaines l'ont été [GBLP16]. D'abord, avant de les énoncer, nous devons observer quelques définitions tirées de [Bog07].

Définition 4.1. Une *tribu* ou σ -*algèbre* \mathcal{A} sur un ensemble X est un ensemble tel que :

1. $\mathcal{A} \neq \emptyset$
2. $\forall B \in \mathcal{A}, B^c \in \mathcal{A}$, où $B \in X$ et B^c est le complémentaire de B dans X
3. si $\forall n \in \mathbb{N}, B_n \in \mathcal{A}$ alors $\bigcup_{n \in \mathbb{N}} B_n \in \mathcal{A}$.

Définition 4.2. Une paire (X, \mathcal{A}) telle que X est un ensemble et \mathcal{A} est une σ -algèbre des sous-ensembles de X est un *espace mesurable*.

Définition 4.3. Soit (X, \mathcal{A}) , un espace mesurable. Une application μ à valeurs dans \mathbb{R}^+ est une *mesure* si :

1. $\mu(\emptyset) = 0$
2. μ est σ -*additive*. C'est-à-dire que si E_1, E_2, \dots est une famille dénombrable de parties de X appartenant à \mathcal{A} et si ces parties sont deux à deux disjointes, alors on a $\mu\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} \mu(E_i)$.

Définition 4.4. Soient μ et ν deux mesures sur un espace mesurable (X, \mathcal{A}) . On dit que ν est *absolument continue* par rapport à μ si et seulement si pour tout A tel que A est une partie de X et $A \in \mathcal{A}$ on a $\mu(A) = 0 \Rightarrow \nu(A) = 0$.

Remarque 4.5. Si X est une partie de \mathbb{R}^2 , ce qui est le cas ici, la mesure μ par défaut est la mesure de Lebesgue qui compte l'aire de surface d'un ensemble [Bog07]. Alors, une autre mesure ν sur X est dite absolument continue si elle est absolument continue par rapport à la mesure de Lebesgue.

Définition 4.6. Soit (X, \mathcal{A}) un espace mesurable et $f : X \rightarrow X$ une fonction mesurable. Une mesure μ sur (X, \mathcal{A}) est dite *invariante* sur f si pour tout A tel que A est une partie de X et $A \in \mathcal{A}$ on a $\mu(f^{-1}(A)) = \mu(A)$. Donc, une mesure est *invariante* si elle est préservée par une fonction.

Après ces quelques définitions, tout est en place pour pouvoir énoncer d'autres conjectures sur le système étudié.

1. Pour $0 < \alpha < \frac{1}{2}$ l'application G conserve une mesure invariante absolument continue.

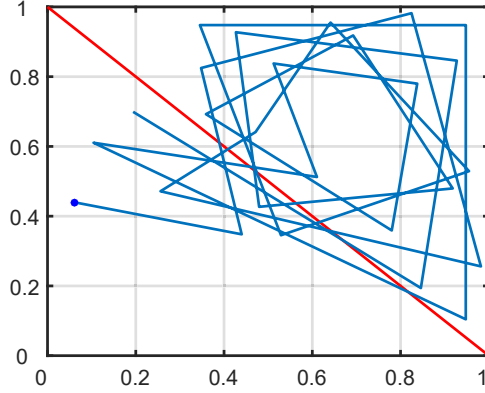
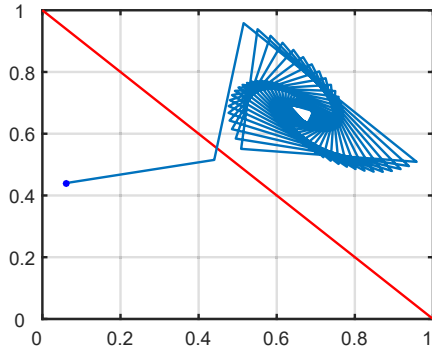
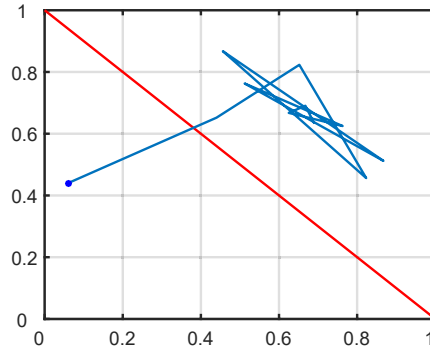


Figure 4 – 25 itérations du point $(0,06;0,44)$ avec $\alpha = 0,3$.

2. Pour $\frac{1}{2} < \alpha < \frac{3}{4}$, le point $(\frac{2}{3}, \frac{2}{3})$ est un attracteur global.



100 itérations avec $\alpha = 0,52$



10 itérations avec $\alpha = 0,7$.

Figure 5 – Itérations du point $(0,06;0,44)$ avec $\frac{1}{2} < \alpha < \frac{3}{4}$.

Cette conjecture a été prouvée.

Les tests effectués avec d'autres points choisis au hasard semblent aussi vouloir montrer que la convergence est plus rapide lorsque α est proche de $\frac{3}{4}$.

3. Pour $\alpha = \frac{3}{4}$, tous les points sur la droite $x + y = \frac{4}{3}$ sont de période deux sauf pour le point fixe $(\frac{2}{3}, \frac{2}{3})$. Tous les autres points $(x, y) \in [0, 1]^2$ tels que $x + y \neq \frac{4}{3}$ et $(x, y) \neq (0, 0)$ sont attirés vers cette droite.

Cette conjecture a aussi été prouvée.

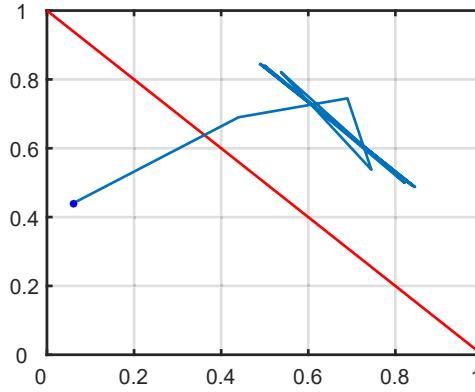
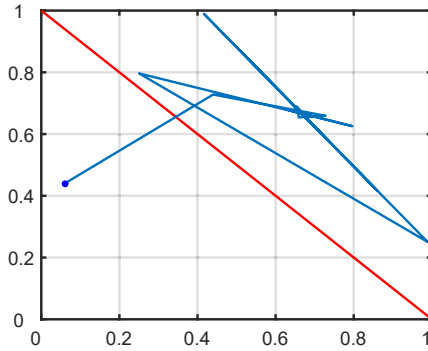
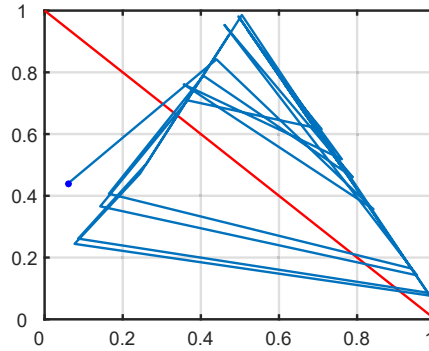


Figure 6 – 10 itérations du point $(0,06; 0,44)$ avec $\alpha = 0,75$.

4. Pour $\frac{3}{4} < \alpha < 1$, l'application G préserve une mesure SRB [You02] qui n'est pas absolument continue. Le concept de mesure appelée SRB, d'après Sinai, Ruelle et Bowen, est une formulation rigoureuse de ce qu'il faut à une mesure invariante pour être significative du point de vue de la physique [Wil08].



20 itérations avec $\alpha = 0,8$.



25 itérations avec $\alpha = 0,95$.

Figure 7 – Itérations du point $(0,06; 0,44)$ avec $\frac{3}{4} < \alpha < 1$.

Références

- [Bog07] Vladimir Igorevich BOGACHEV : *Measure Theory*, volume 1. Springer, 2007.
- [Dev92] Robert L. DEVANEY : *A First course in Chaotic Dynamical Systems*. Westview Press, 1992.
- [Dev03] Robert L. DEVANEY : *An Introduction to Chaotic Dynamical systems*. Westview Press, second edition, 2003.
- [GBLP16] Pawel GÓRA, Abraham BOYARSKY, Zhenyang LI et Harald PROPPE : Statistical and deterministic dynamics of maps with memory. *arXiv [math.DS]*, 2016.
- [Rob95] Clark ROBINSON : *Dynamical Systems: Stability Symbolic Dynamics and Chaos*. CRC Press, 1995.
- [Wil08] Amie WILKINSON : Smooth ergodic theory. *arXiv [math.DS]*, 2008.
- [You02] Lai-Sang YOUNG : What are SRB measures, and which dynamical systems have them? *Journal of Statistical Physics*, 108(516), 2002.

MATHIEU FRAPPIER

DÉPARTEMENT DE MATHÉMATIQUES, UNIVERSITÉ DE SHERBROOKE

Courriel: Mathieu.Frappier@USherbrooke.ca

Classification des représentations indécomposables du carquois de Kronecker

Tanna Sánchez McMillan

Reçu le 2013-07-01 et accepté le 2017-01-13

RÉSUMÉ On considère le carquois de Kronecker \mathbf{K}_2 ayant deux sommets et deux flèches parallèles qui pointent vers la même direction. Soit $M = (E_1, E_2, f, g)$ une représentation de \mathbf{K}_2 , où E_1, E_2 sont des k -espaces vectoriels et f et g des applications k -linéaires allant de E_2 vers E_1 . Dans cet article, on démontre qu'on peut classer les représentations indécomposables de \mathbf{K}_2 en trois classes.

1 Introduction

La théorie de représentations se penche sur l'étude des structures algébriques, telles que les algèbres, par voie de l'analyse de leurs représentations. Dans ce sens, actuellement, les carquois s'avèrent un outil essentiel à la théorie des représentations des algèbres. Plus précisément, à chaque algèbre de dimension finie, sobre et connexe sur un corps k algébriquement clos lui correspond un carquois et vice-versa. De plus, en se servant d'un carquois, on peut visualiser les modules sur l'algèbre associée comme un ensemble de k -espaces vectoriels, un pour chaque sommet du carquois, liés entre eux par des applications linéaires, une pour chaque flèche dans le carquois vérifiant éventuellement des conditions supplémentaires. Ce qu'on vient de décrire est ce qu'on appelle une représentation de carquois.

Toute représentation non nulle d'un carquois fini Q peut, en vertu du théorème de Krull-Schmidt, être décomposée en somme directe finie de représentations indécomposables - soit, des représentations qui ne sont pas isomorphes à la somme directe finie de deux représentations non nulles. De plus, cette décomposition en représentations indécomposables est unique à isomorphisme près. Ainsi, si on comprend toutes les représentations indécomposables d'un carquois Q et les morphismes qui agissent entre eux, on peut comprendre la catégorie des représentations de Q .

En ce qui concerne, en particulier, le carquois de Kronecker, on connaît depuis 1946 [Die46] les classes de ses représentations indécomposables. Plus récemment, on trouve dans [ARO97] cette même classification obtenue en se servant de la

Je tiens à remercier le professeur Ibrahim Assem pour son soutien et ses conseils, ainsi que Juan Carlos Bustamante pour prendre le temps de répondre à mes questions.

14 Classification des représentations indécomposables du carquois de Kronecker
théorie des représentations d'algèbres héréditaires et du carquois d'Auslander-Reiten.

Dans un premier temps, à la Section 2, on introduit les notions de théorie de représentations de carquois nécessaires à la classification. Cette partie ne prétend être qu'une invitation à l'étude plus approfondie de cette théorie. Dans un deuxième temps, à la Section 3, on présente une reformulation de la preuve du théorème de classification des représentations indécomposables de \mathbf{K}_2 décrite dans [Bur86].

2 Représentations de carquois

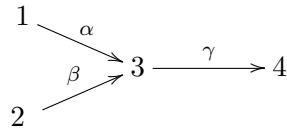
Dans cette section, nous introduisons les concepts de base de théorie de représentations de carquois essentiels à la compréhension de la preuve présentée à la Section 3. Nous référons le lecteur à [Sch14] et [ASS06] pour des plus amples explications à ce sujet. Ici, k représentera un corps algébriquement clos.

2.1 Définitions et préliminaires

Définition 2.1. Un *carquois* est un quadruplet $Q = (Q_0, Q_1, s, b)$ formé de deux ensembles Q_0 et Q_1 , dont les éléments sont appelés respectivement les *sommets* et les *flèches* de Q , et de deux applications $s, b : Q_1 \rightarrow Q_0$ qui associent à chaque flèche $\alpha \in Q_1$ sa *source* $s(\alpha)$, et son *but* $b(\alpha)$, dans Q_0 . On dit que Q est *fini* si Q_0 et Q_1 sont des ensembles finis.

Définition 2.2. Soit Q un carquois et $i, j \in Q_0$. Un *chemin* c de i vers j de longueur l dans Q est une suite $c = (i \mid \alpha_1, \alpha_2, \dots, \alpha_l \mid j)$ avec $\alpha_n \in Q_1$ pour tout $n = 1, 2, \dots, l$ tel que $s(\alpha_1) = i$, $s(\alpha_n) = b(\alpha_{n-1})$ pour tout $n = 2, 3, \dots, l$ et $b(\alpha_l) = j$.

Exemple 2.3. Dans le carquois Q donné par



on a que $Q_0 = \{1, 2, 3, 4\}$ et $Q_1 = \{\alpha, \beta, \gamma\}$. Pour ce qui est des applications source et but, $s(\alpha) = 1$, $b(\alpha) = b(\beta) = s(\gamma) = 3$, $s(\beta) = 2$ et $b(\gamma) = 4$.

Définition 2.4. Une *représentation* $M = (E_i, \phi_\alpha)_{i \in Q_0, \alpha \in Q_1}$ d'un carquois Q est telle que :

1. À chaque sommet i de Q est associé un k -espace vectoriel E_i .
2. À chaque flèche $\alpha : i \rightarrow j$ de Q est associée une application k -linéaire $\phi_\alpha : E_i \rightarrow E_j$.

De plus, une représentation $M = (E_i, \phi_\alpha)_{i \in Q_0, \alpha \in Q_1}$ est dite *nulle* (ou *zéro*) si $E_i = 0$ pour tout $i \in Q_0$.

Exemple 2.5. Soit Q le carquois donné par

$$1 \xrightarrow{\alpha} 2 \xleftarrow{\beta} 3.$$

Donc,

$$M = (k \xrightarrow{1} k \xleftarrow{0} 0)$$

$$\text{et } M' = (k \xrightarrow{\begin{pmatrix} 1 \\ 0 \end{pmatrix}} k^2 \xleftarrow{\begin{pmatrix} 0 \\ 1 \end{pmatrix}} k)$$

sont deux représentations de Q .

Définition 2.6. Soit i un sommet d'un carquois Q .

1. La *représentation simple* $S(i)$ de Q au sommet i est donnée par

$$S(i) = (S(i)_j, \phi_\alpha)_{j \in Q_0, \alpha \in Q_1}$$

où

$$S(i)_j = \begin{cases} k & \text{si } i = j \\ 0 & \text{sinon} \end{cases}$$

2. La *représentation projective* $P(i)$ de Q au sommet i est donnée par $P(i) = (P(i)_j, \phi_\alpha)_{j \in Q_0, \alpha \in Q_1}$, où $P(i)_j$ est le k -espace vectoriel dont la base est l'ensemble de tous les chemins de i à j dans Q et $\phi_\alpha : P(i)_j \rightarrow P(i)_l$ est l'application linéaire définie par la composition de chemins de i à j avec la flèche α qui va de j à l .
3. La *représentation injective* $I(i)$ de Q au sommet i est donnée par $I(i) = (I(i)_j, \phi_\alpha)_{j \in Q_0, \alpha \in Q_1}$, où $I(i)_j$ est le k -espace vectoriel dont la base est l'ensemble de tous les chemins de j à i dans Q et $\phi_\alpha : I(i)_j \rightarrow I(i)_l$ est l'application linéaire définie en enlevant la flèche α des chemins de j à i qui commencent par α et en envoyant vers zéro tous les chemins qui ne commencent pas par α .

Exemple 2.7. Soit Q le carquois de l'Exemple 2.5. Donc,

$$S(1) = (k \longrightarrow 0 \longleftarrow 0)$$

$$P(3) = (0 \longrightarrow k \xleftarrow{1} k)$$

$$I(2) = (k \xrightarrow{1} k \xleftarrow{1} k)$$

Définition 2.8. Soit Q un carquois et soient $M = (E_i, \phi_\alpha)_{i \in Q_0, \alpha \in Q_1}$, $M' = (E'_i, \phi'_\alpha)_{i \in Q_0, \alpha \in Q_1}$ deux représentations de Q . Un *morphisme de représentations* $f : M \longrightarrow M'$ est une famille $(f_i)_{i \in Q_0}$ d'applications linéaires $f_i : E_i \longrightarrow E'_i$ telle que, pour toute flèche $\alpha : i \longrightarrow j$ dans Q_1 , le diagramme

$$\begin{array}{ccc} E_i & \xrightarrow{\phi_\alpha} & E_j \\ f_i \downarrow & & \downarrow f_j \\ E'_i & \xrightarrow{\phi'_\alpha} & E'_j \end{array}$$

commute, c'est-à-dire, $f_j \circ \phi_\alpha(m) = \phi'_\alpha \circ f_i(m)$ pour tout $m \in E_i$.

Exemple 2.9. Soient $M = (k \xrightarrow[\begin{pmatrix} 0 \\ 1 \end{pmatrix}]{\begin{pmatrix} 1 \\ 0 \end{pmatrix}} k^2)$ et $M' = (k \xrightarrow[\begin{pmatrix} 1 \\ 0 \end{pmatrix}]{\begin{pmatrix} 1 \\ 0 \end{pmatrix}} k^2)$ deux représentations du carquois $2 \xrightarrow[\beta]{\alpha} 1$. Ainsi, $f = (f_1, f_2) : M \rightarrow M'$ donné par

$$\begin{array}{ccc} k & \xrightarrow[\begin{pmatrix} 0 \\ 1 \end{pmatrix}]{\begin{pmatrix} 1 \\ 0 \end{pmatrix}} & k^2 \\ f_1=1 \downarrow & & \downarrow f_2 = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \\ k & \xrightarrow[\begin{pmatrix} 1 \\ 0 \end{pmatrix}]{\begin{pmatrix} 1 \\ 0 \end{pmatrix}} & k^2 \end{array}$$

est un morphisme de représentations. En effet, on peut vérifier que

$$\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Définition 2.10. Un morphisme $f = (f_i) : M \rightarrow M'$ de deux représentations d'un carquois Q est *injectif* (*surjectif*) si f_i est injective (surjective) pour tout $i \in Q_0$. De même, f est un *isomorphisme* si f_i est bijective pour tout $i \in Q_0$. D'ailleurs, on note $\text{Hom}(M, M')$ l'ensemble de tous les morphismes allant de M vers M' . Cet ensemble est un k -espace vectoriel pour la somme de morphismes et la multiplication par des scalaires. Un morphisme $f : M \rightarrow M$ est appelé un *endomorphisme* et l'espace vectoriel d'endomorphismes de M est noté $\text{End}(M)$. De plus, pour la composition de morphismes, $\text{End}(M)$ est une k -algèbre associative. En outre, une représentation L d'un carquois Q est appelée *sous-représentation* d'une représentation M s'il existe un morphisme injectif $i : L \hookrightarrow M$.

Définition 2.11. Soient $M = (E_i, \phi_\alpha)_{i \in Q_0, \alpha \in Q_1}$ et $M' = (E'_i, \phi'_\alpha)_{i \in Q_0, \alpha \in Q_1}$ deux représentations de Q . Alors,

$$M \oplus M' = \left(E_i \oplus E'_i, \begin{pmatrix} \phi_\alpha & 0 \\ 0 & \phi'_\alpha \end{pmatrix} \right)_{i \in Q_0, \alpha \in Q_1}$$

est une représentation de Q qui s'appelle *somme directe* de M et M' . Par récurrence, on définit la somme directe d'un nombre fini de représentations $M_1 \oplus M_2 \oplus \dots \oplus M_t = (M_1 \oplus M_2 \oplus \dots \oplus M_{t-1}) \oplus M_t$.

Exemple 2.12. Prenons les représentations M et M' de $1 \xrightarrow{\alpha} 2 \xleftarrow{\beta} 3$ dans l'Exemple 2.5. Alors, la somme directe $M \oplus M'$ est la représentation

$$k \oplus k \xrightarrow{\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}} k \oplus k^2 \xleftarrow{\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}} k$$

qui est isomorphe à

$$k^2 \xrightarrow{\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}} k^3 \xleftarrow{\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}} k$$

Définition 2.13. Une représentation non nulle M de Q est dite *indécomposable* si $M = N \oplus L$, avec N et L aussi des représentations de Q , implique que $N = 0$ ou $L = 0$.

On rappelle au lecteur qu'une algèbre est dite locale lorsque tous ses éléments non inversibles forment un idéal. De façon équivalente, un élément d'une algèbre locale est soit inversible soit nilpotent. On trouve une preuve détaillée du théorème classique suivant dans [Sch14, p.126], [ASS06, p.22].

Théorème 2.14. Soit M une représentation de Q . Alors, M est indécomposable si et seulement si $\text{End}(M)$, vue en tant qu'algèbre, est locale.

2.2 Carquois de Kronecker

Le carquois \mathbf{K}_2 qui est le sujet principal de cet article est le carquois de Kronecker à deux flèches

$$2 \xrightleftharpoons[\beta]{\alpha} 1.$$

On voit que $(\mathbf{K}_2)_0 = \{1, 2\}$, $(\mathbf{K}_2)_1 = \{\alpha, \beta\}$, $s(\alpha) = s(\beta) = 2$ et $b(\alpha) = b(\beta) = 1$. De plus,

$$\begin{aligned} S(1) &= (0 \rightrightarrows k) = P(1), & I(1) &= (k^2 \rightrightarrows k), \\ S(2) &= (k \rightrightarrows 0) = I(2), & P(2) &= (k \rightrightarrows k^2) \end{aligned}$$

sont ses représentations simples, projectives et injectives.

On présente maintenant quelques représentations de \mathbf{K}_2 pertinentes pour la preuve du théorème de classification.

Pour $n \geq 1$, on note \bar{P}^n la représentation $k^n \xrightleftharpoons[g]{f} k^{n+1}$ appelée *postprojective*. Dans celle-ci, l'espace vectoriel E_2 est de dimension n et l'espace vectoriel E_1 de dimension $n+1$. On fixe une base $\{u_1, \dots, u_n\}$ de E_2 et une base $\{v_1, \dots, v_{n+1}\}$ de E_1 et on définit f et g telles que $f(u_i) = v_i$, $g(u_i) = v_{i+1}$. Remarquons que, puisque l'image par f ou g d'une base de E_2 est un ensemble linéairement indépendant de E_1 , ces applications sont injectives. On définit ensuite, pour $n \geq 1$, la représentation dite *préinjective* $k^{n+1} \xrightleftharpoons[g]{f} k^n$ qu'on note \bar{Q}^n . Pour cette représentation, f et g sont telles que $f(u_1) = g(u_{n+1}) = 0$, $f(u_i) = v_{i-1}$ pour tout $i \in \{2, \dots, n+1\}$ et $g(u_i) = v_i$ pour tout $i \in \{1, \dots, n\}$.

Proposition 2.15. *Les représentations \bar{P}^n et \bar{Q}^n de K_2 sont indécomposables.*

Démonstration. On montre que \bar{P}^n est indécomposable, car la preuve pour \bar{Q}^n est semblable. En vertu du Théorème 2.14, il suffit de montrer que $\text{End}(\bar{P}^n)$ est locale. Soit $h = (h_1, h_2)$ un endomorphisme de \bar{P}^n où $h_1 : E_2 \rightarrow E_2$ et $h_2 : E_1 \rightarrow E_1$ sont telles que $h_1(u_i) = u_{i+k}$ et $h_2(v_j) = v_{j+k}$ pour $k \in \{0, 1, \dots, n+1\}$. On peut vérifier que $h_2 \circ f = f \circ h_1$ et $h_2 \circ g = g \circ h_1$. Les applications linéaires h_1 et h_2 peuvent être représentées par une matrice carrée d'ordre n ayant des zéros partout sauf dans l'une de ses diagonales où les éléments sont tous des 1. Ainsi, si $k = 0$, dite matrice est la matrice identité et alors h est un isomorphisme. Si $k \neq 0$, la matrice est nilpotente. Donc, $\text{End}(\bar{P}^n)$ est une algèbre locale. \square

Fixons, maintenant, un entier $n \geq 1$ et un $\lambda \in k$. Considérons le $n \times n$ bloc de Jordan $J_n(\lambda) = \begin{pmatrix} \lambda & 0 & 0 & \dots & 0 \\ 1 & \lambda & 0 & \dots & 0 \\ 0 & 1 & \lambda & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda \end{pmatrix}$. On note $H_\lambda^n = (k^n \xrightarrow[J_n(\lambda)]{I_n} k^n)$

et $K_\lambda^n = (k^n \xrightarrow[I_n]{J_n(\lambda)} k^n)$ deux représentations de K_2 où I_n est la matrice identité. Les représentations H_λ^n et K_λ^n s'appellent *régulières*. On démontre que ces représentations sont indécomposables.

Proposition 2.16. *Les représentations H_λ^n et K_λ^n de K_2 , pour tous $n \geq 1$ et $\lambda \in k$, sont indécomposables.*

Démonstration. Montrons que H_λ^n est indécomposable. Soit $h = (A, B)$ un endomorphisme de H_λ^n où A et B sont des matrices carrées d'ordre n à coefficients dans k .

$$\begin{array}{ccc} k & \xrightarrow[J_n(\lambda)]{I_n} & k^2 \\ A \downarrow & & \downarrow B \\ k & \xrightarrow[J_n(\lambda)]{I_n} & k^2 \end{array}$$

On veut que $A = B$ et que $B J_n(\lambda) = J_n(\lambda) A$. Soit $A = B = [a_{i,j}]$ où $a_{i,j} \in k$ (pour tout i, j). On vérifie sans peine que l'égalité des produits matriciels $B J_n(\lambda)$ et $J_n(\lambda) A$ entraîne que A est de la forme

$$A = \begin{pmatrix} a_{1,1} & 0 & 0 & \dots & 0 \\ a_{2,1} & a_{1,1} & 0 & \dots & 0 \\ a_{3,1} & a_{2,1} & a_{1,1} & \dots & 0 \\ \ddots & \ddots & \ddots & \ddots & \vdots \\ a_{n,1} & a_{n-1,1} & \dots & a_{2,1} & a_{1,1} \end{pmatrix}.$$

On a donc que $\text{End}(H_\lambda^n)$ est isomorphe à l'anneau R des matrices carrées d'ordre n de la forme ci-dessus. Ainsi, A est non inversible si et seulement si $\det A = 0$, si et seulement si $a_{1,1} = 0$ et on peut vérifier que

$$I = \left\{ \left(\begin{pmatrix} 0 & 0 & 0 & \dots & 0 \\ a_{2,1} & 0 & 0 & \dots & 0 \\ a_{3,1} & a_{2,1} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n-1,1} & \dots & a_{2,1} & 0 \end{pmatrix} \middle| a_{i,j} \in k \right\} \subseteq R$$

est un idéal de R . Ainsi, R est local et donc $\text{End}(H_\lambda^n)$ l'est aussi. On prouve de même que K_λ^n est aussi une représentation indécomposable de \mathbf{K}_2 . \square

3 Classification des représentations indécomposables

Cette section représente le coeur de l'article. Nous y présentons la classification des représentations indécomposables de \mathbf{K}_2 . Plus précisément, on démontre que toute représentation indécomposable de \mathbf{K}_2 est soit postprojective, soit préinjective, soit régulière. En vertu du théorème de Krull-Schmidt (dont la preuve se trouve, par exemple, dans [Sch14, p.11]), une représentation se décompose de manière unique en une somme directe de représentations indécomposables, ceci à isomorphisme près et à l'ordre des facteurs près. Dans le but d'obtenir la classification, nous dresserons la liste complète, à isomorphisme près, des représentations indécomposables de \mathbf{K}_2 . Pour y parvenir, on vise d'abord à mettre en évidence certains sous-espaces de E_1 et de E_2 .

Occupons-nous, d'abord, des cas triviaux.

Lemme 3.1. *Soit $M = (E_2 \xrightarrow[f]{g} E_1)$ une représentation indécomposable de \mathbf{K}_2 . Si M est non simple, on a que*

1. $\text{Ker } f \cap \text{Ker } g = 0$
2. $\text{Im } f + \text{Im } g = E_1$

Démonstration. 1. Posons $F = \text{Ker } f \cap \text{Ker } g \neq 0$. Il existe alors un facteur direct de M de la forme $F \xrightarrow[f]{g} 0$. Puisque M est indécomposable, alors

$$M = (F \xrightarrow[f]{g} 0) \text{ et } \dim F = 1. \text{ Par conséquent, } M \cong (k \xrightarrow{\quad} 0) \cong S(2).$$

2. Supposons $\text{Im } f + \text{Im } g \subsetneq E_1$ et soit F un sous-espace vectoriel de E_1 tel que $E_1 = F \oplus (\text{Im } f + \text{Im } g)$. Il existe alors un facteur direct de M de la forme $(0 \xrightarrow[f]{g} F)$ et donc $M \cong (0 \xrightarrow{\quad} F) \cong S(1)$.

\square

Remarquons qu'une représentation $E_2 \xrightarrow[f]{g} E_1$ est indécomposable si et seulement si la représentation duale $E_1^* \xrightarrow[g^*]{f^*} E_2^*$ est indécomposable. Ainsi, on peut, sans perte de généralité, limiter l'étude des représentations indécomposables à $M = (E_2 \xrightarrow[f]{g} E_1)$ avec $\dim E_2 \leq \dim E_1$.

Soit M une telle représentation indécomposable non simple de \mathbf{K}_2 . Supposons $\dim E_2 = n$ et $\dim E_1 = n + m$ avec $m \in \mathbb{N}$. Au vu du Lemme 3.1, on peut supposer dès ici, sans perte de généralité que $\text{Ker } f \cap \text{Ker } g = 0$ et $\text{Im } f + \text{Im } g = E_1$. Notons $d = \dim(\text{Ker } f)$ et $d' = \dim(\text{Ker } g)$. Finalement, soit $W = \text{Im } f \cap \text{Im } g$ et $V = f^{-1}(W) \cap g^{-1}(W)$ des sous-espaces vectoriels de E_1 et E_2 respectivement.

Notons que M a une sous-représentation de la forme $V \xrightarrow[g_V]{f_V} W$ où f_V et g_V sont les restrictions de f et g à V . Le lemme qui suit fournit davantage d'information sur cette sous-représentation.

Lemme 3.2. *On a que*

$$a) \dim W = n - d - d' - m$$

$$b) \dim V \geq n - d - d' - 2m$$

Démonstration.

a) On a supposé que $E_1 = \text{Im } f + \text{Im } g$. Utilisant la formule de Grassmann, on a que

$$\dim E_1 = (\dim E_2 - \dim(\text{Ker } f)) + (\dim E_2 - \dim(\text{Ker } g)) - \dim W$$

alors,

$$n + m = (n - d) + (n - d') - \dim W$$

et donc,

$$\dim W = n - d + n - d' - n - m = n - d - d' - m.$$

b) On a que $ff^{-1}(W) = W \cap \text{Im } f = W$ donc $f : f^{-1}(W) \rightarrow W$ est surjective et

$$\begin{aligned} \dim f^{-1}(W) &= \dim W + \dim(\text{Ker } f) \\ &= (n - d - d' - m) + d \\ &= n - d' - m. \end{aligned}$$

De même, $\dim g^{-1}(W) = n - d - m$. Ainsi,

$$\begin{aligned} \dim f^{-1}(W) + \dim g^{-1}(W) &= n - d' - m + n - d - m \\ &= 2n - d' - d - 2m. \end{aligned}$$

Finalement, $f^{-1}(W) + g^{-1}(W) \subseteq E_2$ et donc $\dim(f^{-1}(W) + g^{-1}(W)) \leq n$. Alors $\dim f^{-1}(W) + \dim g^{-1}(W) - \dim(f^{-1}(W) \cap g^{-1}(W)) \leq n$, c'est-à-dire $(2n - d' - d - 2m) - \dim V \leq n$, ce qui entraîne que $\dim V \geq n - d - d' - 2m$. \square

Continuons l'identification de sous-espaces de E_2 . Soient

- i) $K = \text{Ker } f \cap V$ et $K' = \text{Ker } g \cap V$ de dimensions respectives k et k' et
- ii) L et L' deux sous-espaces vectoriels de E_2 tel que $\text{Ker } f = L \oplus K$ et $\text{Ker } g = L' \oplus K'$ avec $\dim L = l$ et $\dim L' = l'$.

Ainsi, $d = k + l$ et $d' = k' + l'$.

Du fait que $L \cap V = L \cap \text{Ker } f \cap V = L \cap K = 0$, on a que $V \oplus L$ est un sous-espace de $f^{-1}(W)$. Nommons H le sous-espace de $f^{-1}(W)$, tel que $f^{-1}(W) = V \oplus L \oplus H$ et $r \geq 0$ tel que $\dim V = n - d - d' - 2m + r$. Conséquentment,

$$\begin{aligned} \dim H &= \dim f^{-1}(W) - \dim L - \dim V \\ &= (n - d' - m) - (d - k) - (n - d - d' - 2m + r) \\ &= k + m - r. \end{aligned}$$

On définit de même H' tel que $g^{-1}(W) = V \oplus L' \oplus H'$ et donc $\dim H' = k' + m - r$.

Lemme 3.3. *La somme $(V \oplus L \oplus H) + (L' \oplus H')$ est directe et sa dimension est $n - r$.*

Démonstration. Prenons $x \in (V \oplus L \oplus H) \cap (L' \oplus H')$, alors $x \in V + L + H = f^{-1}(W)$ et $x \in L' + H' \subseteq g^{-1}(W)$ et donc $x \in V$. Mais $V \cap (L' \oplus H') = 0$ et ainsi $x = 0$, ce qui démontre que $(V \oplus L \oplus H) + (L' \oplus H')$ est une somme directe. En outre,

$$\begin{aligned} \dim(V \oplus L \oplus H \oplus L' \oplus H') &= \dim V + \dim L + \dim H + \dim L' + \dim H' \\ &= (n - d - d' - 2m + r) + (d - k) + (d' - k') + (k + m - r) + (k' + m - r) \\ &= n - r. \end{aligned}$$

□

Maintenant, soit X un sous-espace vectoriel de E_2 de dimension r tel que $E_2 = V \oplus L \oplus H \oplus L' \oplus H' \oplus X$, et $Y = f(X) + g(X) \subseteq E_1$. De plus, comme $L \subseteq \text{Ker } f$, on a que $f(f^{-1}(W)) = f(H) + f(V)$, car $f^{-1}(W) = V \oplus L \oplus H$, et puisque, comme démontré à la partie b) du Lemme 3.2, $ff^{-1}(W) = W$, on a $f(H) + f(V) = W$. De même, $g(V) + g(H') = W$. Ainsi,

$$\begin{aligned} E_1 &= \text{Im } f + \text{Im } g \\ &= W + g(L) + f(L') + g(H) + f(H') + Y \end{aligned}$$

Soit s la somme des dimensions de ces derniers sous-espaces. Ainsi, $\dim E_1 \leq s$, c'est-à-dire, $n + m \leq s$. Par ailleurs,

$$s \leq (n - d - d' - m) + (d - k) + (d' - k') + (k + m - r) + (k' + m - r) + 2r = n + m.$$

Par conséquent, $s = n + m$ et on a, premièrement, que $\dim Y = 2r$, deuxièmement, que

$$\begin{aligned} f(L') &\cong L' \\ f(H') &\cong H' \\ g(L) &\cong L \\ g(H) &\cong H \end{aligned}$$

sont des sous-espaces isomorphes et, finalement, que $E_1 = W \oplus g(L) \oplus f(L') \oplus g(H) \oplus f(H') \oplus Y$. On a ainsi réussi à décomposer E_1 et E_2 en somme directe de sous-espaces vectoriels.

$$\begin{aligned} E_1 &= W \oplus g(L) \oplus f(L') \oplus g(H) \oplus f(H') \oplus Y \\ E_2 &= V \oplus L \oplus L' \oplus H \oplus H' \oplus X \end{aligned}$$

Corollaire 3.4. *Avec les notations ci-dessus, une représentation*

$M = (E_2 \xrightleftharpoons[g]{f} E_1)$ *est une somme directe de quatre sous-représentations :*

- a) $L \xrightleftharpoons{\quad} g(L)$
- b) $L' \xrightleftharpoons{\quad} f(L')$
- c) $X \xrightleftharpoons{\quad} Y$
- d) $V \oplus H \oplus H' \xrightleftharpoons{\quad} W \oplus f(H') \oplus g(H)$

où les applications sont les restrictions de f et g aux sous-espaces correspondants.

Lorsque M est indécomposable, M est isomorphe à une de ces quatre représentations. Dans le lemme suivant, on considère les trois cas les plus simples.

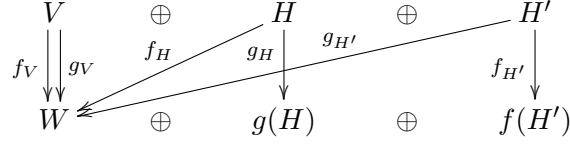
Lemme 3.5. *On a que*

- a) Si $M = (L \xrightleftharpoons{\quad} g(L))$ alors $M \cong K_0^1$
- b) Si $M = (L' \xrightleftharpoons{\quad} f(L'))$ alors $M \cong H_0^1$
- c) Si $M = (X \xrightleftharpoons{\quad} Y)$ alors $M \cong P(2)$

Démonstration.

a) Supposons que $M = (L \xrightleftharpoons{\quad} g(L))$. On sait que la restriction de f à L est zéro et, ainsi, M est indécomposable si et seulement si la dimension de L est 1.

b) Se démontre comme la partie a).

**Figure 1** – Applications de la représentation M

c) On a $\dim Y = 2r = 2 \dim X$. Par conséquent, $Y = f(X) \oplus g(X)$. Alors, M est indécomposable si et seulement si $\dim X = 1$ et $\dimens Y = 2$. Ainsi,

$$M = (k \xrightarrow{\begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}} k^2) \cong P(2) . \quad \square$$

Il ne reste à s'occuper que du dernier cas, soit

$$M = (V \oplus H \oplus H' \xrightarrow{\quad} W \oplus f(H') \oplus g(H)).$$

Remarquons que, dans ce cas, les sous-espaces X, Y, L et L' sont tous nuls et alors $d = k$, $d' = k'$ et $r = 0$. On a donc que

1. $\dim V = n - d - d' - 2m$
2. $\dim W = n - d - d' - m$
3. $\dim H = k + m = d + m$
4. $\dim H' = k' + m = d' + m$

De plus, puisque $\text{Ker } f \subseteq f^{-1}(W) = V \oplus H$ et $K \cap H \subseteq V \cap H = 0$, alors $\text{Ker } f \subseteq V$. De même, $\text{Ker } g \subseteq V$. Le diagramme dans la Figure 1 montre comment agissent les applications dans M .

Lemme 3.6. *Soit M la représentation décrite dans la Figure 1. La sous-représentation $V \xrightarrow{\quad} W$ est indécomposable.*

Démonstration. On veut montrer qu'une décomposition non triviale

$$(V \xrightarrow{\quad} W) \cong (V_1 \xrightarrow{\quad} W_1) \oplus (V_2 \xrightarrow{\quad} W_2)$$

entraîne une décomposition non triviale de M .

Observons d'abord que la somme $f(V) + f(H)$ est directe. En effet, si $f(v) = f(h)$ avec $v \in V$, $h \in H$, on a $h - v \in \text{Ker } f \subseteq V$. Donc $h \in V$ et par conséquent $h = 0$. On peut donc écrire $W = f(V) \oplus f(H) \oplus W^*$ avec $W^* \subseteq \text{Im } g$. En outre, $H \cap \text{Ker } f = 0$ entraîne que H est isomorphe à $f(H)$.

Par conséquent, la décomposition $W = W_1 \oplus W_2$ donne les décompositions $W_1 = f(V_1) \oplus W'_1$, $W_2 = f(V_2) \oplus W'_2$ avec $f(H) \subseteq W'_1 \oplus W'_2$. Puisque $H \cong f(H)$, il existe deux sous-espaces U_1, U_2 de H tels que $H = U_1 \oplus U_2$ et on a $f(U_1) \subseteq W'_1 \subseteq W_1$ et $f(U_2) \subseteq W'_2 \subseteq W_2$.

Soient W'_1, W'_2 tels que $W_1 = f(V_1) \oplus W'_1$ et $W_2 = f(V_2) \oplus W'_2$. On choisit une décomposition $H = U_1 \oplus U_2$ telle que $f(U_1) = W'_1 \subseteq W_1$ et $f(U_2) = W'_2 \subseteq W_2$. On aura, de même, $H' = U'_1 \oplus U'_2$ telle que $g(U'_1) \subseteq W_1$ et $g(U'_2) \subseteq W_2$. On obtient, ainsi, la décomposition non triviale en sommes directes

$$\begin{aligned} M \cong & ((V_1 \oplus U_1 \oplus U'_1) \rightrightarrows (W_1 \oplus g(U_1) \oplus f(U'_1))) \\ & \oplus ((V_2 \oplus U_2 \oplus U'_2) \rightrightarrows (W_2 \oplus g(U_2) \oplus f(U'_2))) \end{aligned}$$

□

On est maintenant en mesure de compléter la classification. Toujours en lien avec la représentation $M = (V \oplus H \oplus H' \rightrightarrows W \oplus f(H') \oplus g(H))$, considérons trois cas :

Premier cas Supposons que $m = 0$. Dans ce cas, $\dim E_2 = \dim E_1 = n$, $\dim V = \dim W = n - d - d'$, $\dim H = d$ et $\dim H' = d'$.

Lemme 3.7. *Si $m = 0$, alors au moins l'une de f et g est un isomorphisme.*

Démonstration. Procédons par récurrence sur n . Si $n = 1$, l'énoncé est trivial. Supposons alors que $n > 1$.

- Si $d = d' = 0$, alors $\text{Ker } f = \text{Ker } g = 0$. Puisque f et g sont injectives et $\dim E_2 = \dim E_1$, ce sont des isomorphismes.
- Si $d + d' > 0$, en vertu du Lemme 3.6, $V \rightrightarrows W$ est indécomposable et on a $\dim V = \dim W < n$. En vertu de l'hypothèse de récurrence, l'une des restrictions de f et g à V est un isomorphisme. Supposons que f_V est un isomorphisme, alors $f^{-1}(W) = V$ et donc $H = 0$. Dans ces conditions, $f : V \oplus H' \rightarrow W \oplus f(H')$ est un isomorphisme.

□

Corollaire 3.8. *Si $m = 0$, alors $M \cong H_\lambda^n$ avec $\lambda \in k$ ou $M \cong K_0^n$.*

Démonstration. En vertu du Lemme 3.7, M est isomorphe à une représentation de la forme $E \xrightarrow[g]{1} E$ ou $E \xrightarrow[1]{f} E$. Dans le premier cas, munissons E d'une structure d'un $k[t]$ -module en posant $t.x = g(x)$ pour $x \in E$. De plus, E est un $k[t]$ -module indécomposable (sinon, on pourrait décomposer E en sous-espaces stables par g , ce qui contredirait le fait que M est indécomposable) et de type fini. Le théorème fondamental pour les modules sur un anneau principal [AL09, p.520] entraîne qu'il existe $p \in k[t]$ irréductible tel que $E \cong k[t]/\langle p^s \rangle$ pour un s quelconque. Puisque

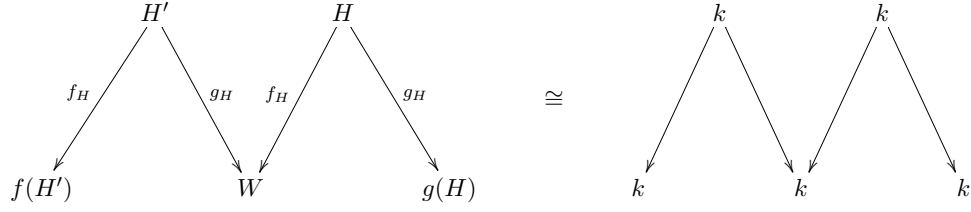


Figure 2 – Si $m = 1$ et $n = 2$, alors $M \cong \bar{P}^2$

k est algébriquement clos, on peut choisir une base appropriée de E telle que g puisse s'écrire dans la forme normale de Jordan comme un bloc $J_n(\lambda)$ pour $\lambda \in k$. Ainsi, $M \cong H_\lambda^n$. De façon analogue, on démontre que si $M = (E \xrightarrow[f]{f} E)$, alors $M \cong K_\lambda^n$. Si $\lambda \neq 0$, alors $J_n(\lambda)$ est inversible et donc f l'est aussi. Ainsi, $M = (E \xrightarrow[f^{-1}]{1} E)$, ce que nous ramène au premier cas. \square

Deuxième cas Supposons que $m = 1$. On a alors que $\dim E_2 = n$, $\dim E_1 = n + 1$, $\dim V = n - d - d' - 2$, $\dim W = n - d - d' - 1$, $\dim H = d + 1$ et $\dim H' = d' + 1$.

Lemme 3.9. *Si, $m = 1$, alors $M \cong \bar{P}^n$.*

Démonstration. Procédons par récurrence sur n . Si $n = 1$, et que f et g sont non nulles, on a que $M \cong (K \xrightarrow[g]{f} K^2) = P(2) \cong \bar{P}^1$. Supposons, maintenant, que $n = 2$, donc $\dim V = n - d - d' - 2$ et, ainsi, $d = d' = 0$ et $V = 0$. Alors, $\dim W = \dim H = \dim H' = 1$. La représentation M peut donc s'écrire sous la forme présentée dans la Figure 2. Ainsi, $M \cong \bar{P}^2$.

Finalement, supposons que $n > 2$. En vertu du Lemme 3.6, $V \xrightarrow{\quad} W$ est indécomposable. Notons que cela entraîne immédiatement que les applications f et g sont non nulles. Par ailleurs, puisque $\dim V > \dim W$ on a que $V \xrightarrow{\quad} W$ n'est ni isomorphe à K_λ^n ni à H_λ^n , pour un λ . Soit $t = \dim V = n - d - d' - 2$ et $\dim W = t + 1$. Puisque $t < n$, en vertu de l'hypothèse de récurrence, on a $V \xrightarrow{\quad} W \cong \bar{P}^t$. En particulier, chacune des restrictions f_V et g_V sont injectives. Par ailleurs, on sait que chacune de $f : H' \cong f(H')$ et $g : H \cong g(H)$ est un isomorphisme d'espaces vectoriels de dimension 1. En outre, $g(H') \neq 0$, car sinon la représentation $k \xrightarrow[0]{1} k$ serait un facteur direct de M . Puisque H' est de dimension 1, cela implique que $g : H' \cong g(H')$ est un isomorphisme entre sous-espaces de dimension 1. De même, $f : H \cong f(H)$ l'est aussi. Conséquemment, M peut s'écrire sous la forme présentée à la Figure 3. \square

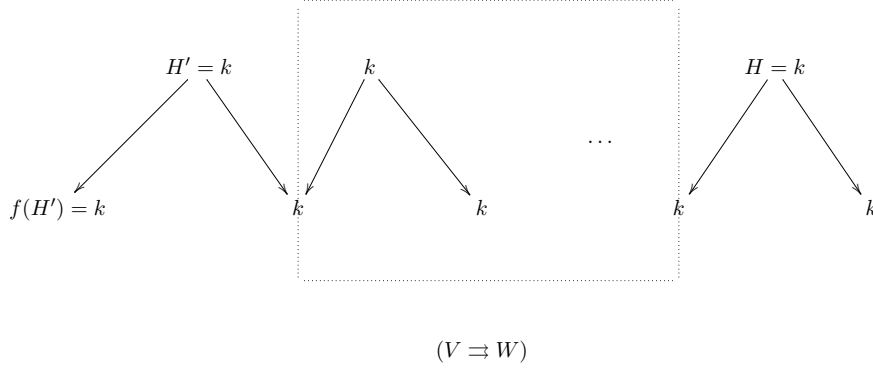


Figure 3 – Si $m = 1$ et $n > 2$, alors $M \cong \bar{P}^n$

Troisième cas Considérons la situation où $m \geq 2$.

Lemme 3.10. *Si $m \geq 2$, alors M n'est plus une représentation indécomposable.*

Démonstration. On procède par récurrence sur n . Si $n = 1$,

$$M \cong (k \xrightleftharpoons[g]{f} k^3) \cong S(1) \oplus \bar{P}^1.$$

Si $n > 1$, considérons la sous-représentation $V \rightrightarrows W$. Donc, $\dim V = n - d - d' - 2m$ et $\dim W = n - d - d' - m = m + \dim V$. Par l'hypothèse de récurrence, $V \rightrightarrows W$ est décomposable et, en vertu du Lemme 3.6, M est aussi décomposable. \square

On a réussi à démontrer le Théorème de classification.

Théorème 3.11 (Théorème de classification). *Soit $M = (E_2 \rightrightarrows E_1)$ une représentation indécomposable de \mathbf{K}_2 . Soit $n = \dim E_2$.*

a) *Si $n < \dim E_1$, alors $M \cong \bar{P}^n$*

b) *Si $n > \dim E_1$, alors $M \cong \bar{Q}^n$*

c) *Si $n = \dim E_1$, alors $M \cong H_\lambda^n$ où $\lambda \in k$ ou $M \cong K_0^n$* \square

On réfère le lecteur à [Bur86] et [Rye13] : le premier a été l'inspiration principal de cet article et le dernier présente une preuve alternative à celle décrite ici, ainsi qu'une application de la décomposition des représentations de \mathbf{K}_2 à la résolution des systèmes d'équations différentielles linéaires $Ax = B\dot{x}$.

Références

- [AL09] Ibrahim ASSEM et Pierre Yves LEDUC : *Cours d'algèbre : groupes, anneaux, modules et corps*. Presses internationales Polytechnique, Québec, 2009.
- [ARO97] Maurice AUSLANDER, Idun REITEN et Sverre O.SMALO : *Representation Theory of Artin Algebras*. Cambridge University Press, Cambridge, 1997.
- [ASS06] Ibrahim ASSEM, Daniel SIMSON et Andrzej SKOWRONSKI : *Elements of the Representation Theory of Associative Algebras*. Cambridge University Press, Cambridge, 2006.
- [Bur86] Pierre-François BURGERMEISTER : Classification des représentations de la double flèche. *L'Enseignement mathématique*, 32:199–210, 1986.
- [Die46] Jean DIEUDONNÉ : Sur la réduction canonique des couples de matrices. *Bulletin de la Société Mathématique de France*, 74:130–146, 1946.
- [Rye13] August Benjamin RYE : The 2-Kronecker Quiver and Systems of Linear Differential Equations. Mémoire de D.E.A., Norwegian University of Science and Technology, 2013.
- [Sch14] Ralf SCHIFFLER : *Quiver Representations*. Springer International Publishing, Suisse, 2014.

TANNA SÁNCHEZ MCMILLAN
DÉPARTEMENT DE MATHÉMATIQUES, UNIVERSITÉ DE SHERBROOKE
Courriel: tanna.sanchez@USherbrooke.ca

Géométrie hyperbolique : le demi-plan de Poincaré

Kevin Thouin

Reçu le 2017-01-11 et accepté le 2017-05-25

RÉSUMÉ Cet article traite de la géométrie hyperbolique de dimension 2. Il sera question du modèle du demi-plan de Poincaré. Après avoir correctement défini et décrit les droites hyperboliques du demi-plan de Poincaré, nous trouverons un groupe de transformations envoyant ces droites hyperboliques vers d'autres droites hyperboliques. Ensuite, on définira une métrique, qui nous permettra de définir les notions de longueur d'un chemin et de distance entre deux points de demi-plan de Poincaré. Finalement, à l'aide d'une courte introduction à la trigonométrie en géométrie hyperbolique, nous montrerons un résultat qui montre une particularité de la géométrie hyperbolique par rapport à la géométrie euclidienne usuelle.

1 Introduction

Pour introduire la géométrie hyperbolique, il est naturel de parler du cinquième postulat d'Euclide. Celui-ci stipule que si nous avons une droite et un point à l'extérieur de cette droite, alors il existe une seule droite parallèle à la première et passant par le point. Dans la géométrie hyperbolique, au contraire, il existe une infinité de telles droites. Nous ne parlerons, par contre, pas de parallélisme dans cet article.

Il existe plusieurs façons de représenter cette géométrie. Dans cet article, nous la représenterons par le demi-plan de Poincaré. Le livre [And05], qui sera notre principale référence, présente d'autres de ces modèles sans être exhaustif. En particulier, il généralise le demi-plan de Poincaré à tout autre demi-plan ou disque.

2 Définitions et généralités

Cette section introduit le demi-plan de Poincaré, un exemple de géométrie hyperbolique. On discutera aussi de la sphère de Riemann.

Définition 2.1. On définit le demi-plan de Poincaré

$$\mathcal{H} = \{z \in \mathbb{C} \mid \text{Im}(z) > 0\}.$$

J'aimerais remercier Virginie Charette pour ses commentaires lors de la rédaction de cet article.

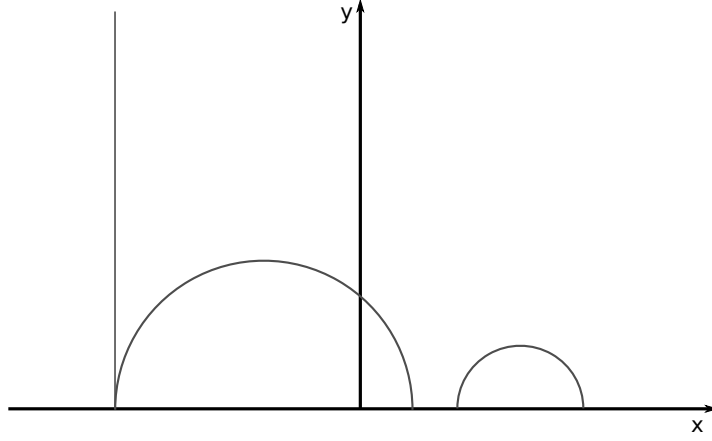


Figure 1 – Trois exemples de droites hyperboliques. Ils ne contiennent pas de points sur l’axe des x comme l’indique la définition.

C’est dans cet ensemble que nous allons définir la géométrie hyperbolique dans cet article. On peut y définir des figures analogues aux droites euclidiennes.

Définition 2.2. On appelle droite hyperbolique tout sous-ensemble de \mathcal{H} satisfaisant une des deux conditions suivantes :

- L’intersection de \mathcal{H} et d’une droite euclidienne verticale dans \mathbb{C} ;
- L’intersection de \mathcal{H} et d’un cercle euclidien centré sur un nombre réel dans \mathbb{C} .

La figure 1 donne trois exemples de droites hyperboliques. La définition d’une droite peut sembler étrange, car elle peut faire intervenir deux objets différents. Toutefois, il est possible d’unifier ces deux concepts comme nous le verrons.

Définition 2.3. La sphère de Riemann est $\hat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$.

La sphère de Riemann est donc le plan complexe muni d’un point additionnel noté ∞ .

Il est possible de définir des notions d’ensembles ouverts et fermés ainsi que de continuité d’une fonction de $\hat{\mathbb{C}}$ vers $\hat{\mathbb{C}}$. Toutefois, nous ne ferons pas ces constructions dans cet article.

Définition 2.4. On appelle cercle généralisé tout sous-ensemble de $\hat{\mathbb{C}}$ qui est soit un cercle dans \mathbb{C} , soit $l \cup \{\infty\}$ où l est une droite euclidienne dans \mathbb{C} .

Proposition 2.5. On peut obtenir une paramétrisation unifiée pour les cercles généralisés. Ils vérifient :

$$\alpha z \bar{z} + \beta z + \bar{\beta} \bar{z} + \gamma = 0 \quad \text{où } \alpha, \gamma \in \mathbb{R} \text{ et } \beta \in \mathbb{C}.$$

De plus, il s’agit d’une droite si, et seulement si $\alpha = 0$.

Démonstration.

- Pour les droites on pose $x = \operatorname{Re}(z) = \frac{1}{2}(z + \bar{z})$ et $y = \operatorname{Im}(z) = \frac{i}{2}(\bar{z} - z)$ et on remplace dans l'équation $ax + by + c = 0$ pour obtenir :

$$\frac{a - bi}{2}z + \frac{a + bi}{2}\bar{z} + c = 0,$$

qui est de la forme désirée avec $\alpha = 0$, $\beta = \frac{a-bi}{2}$ et $\gamma = c$.

- Pour les cercles, on développe l'équation $|z - z_0|^2 - r^2 = 0$ pour obtenir :

$$\bar{z}z - \bar{z}_0z - z_0\bar{z} + (|z_0|^2 - r^2) = 0,$$

qui est de la forme désirée avec $\alpha = 1$, $\beta = -\bar{z}_0$ et $\gamma = |z_0|^2 - r^2$

□

Remarque 2.6. On peut considérer que ∞ est une solution dans le cas où on a une droite euclidienne, c'est-à-dire si $\alpha = 0$. On ne fournira pas de justifications supplémentaires dans cet article, le livre [And05] en donne plus.

Maintenant que nous avons unifié les notions de droite et de cercle, on peut énoncer la définition suivante.

Définition 2.7. Soit d une droite hyperbolique dans \mathcal{H} , on appelle les deux points de l'intersection du cercle généralisé $\widehat{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$ et du cercle contenant d , les points à l'infini de d .

La définition suivante nous servira plus tard lorsque nous trouverons un groupe d'isométrie pour cette géométrie qui envoie des droites hyperboliques sur des droites hyperboliques.

Définition 2.8. Un groupe G agit transitivement sur un ensemble E si pour chaque $x, y \in E$, il existe $g \in G$ tel que $g(x) = y$.

Il est toutefois possible d'avoir une condition plus facile à vérifier.

Lemme 2.9. Soit G un groupe qui agit sur un ensemble E et $x_0 \in E$. Supposons que pour tout $y \in E$, il existe $g \in G$ tel que $g(y) = x_0$. Alors, G agit transitivement sur E .

Démonstration. Soit $x, y \in E$ et $g_1, g_2 \in G$ tel que $g_1(x) = g_2(y) = x_0$. Alors $g_2^{-1} \circ g_1(x) = y$. □

3 Transformations de Möbius

Dans cette section, on définit les transformations de Möbius. Puis, on prouve qu'ils forment un groupe qu'il est possible de représenter par des matrices 2×2 . On trouve finalement un sous-groupe préservant le demi-plan de Poincaré et envoyant les droites hyperboliques sur des droites hyperboliques.

Définition 3.1. Une transformation de Möbius est une fonction de la forme :

$$\begin{aligned} m : \widehat{\mathbb{C}} &\rightarrow \widehat{\mathbb{C}} \\ z &\mapsto \frac{az+b}{cz+d} \quad \text{où } a, b, c, d \in \mathbb{C} \text{ et } ad - bc \neq 0. \end{aligned}$$

On gère l'infini comme suit :

- Si $z = \infty$, $m(z) = \frac{a}{c}$ ou $m(z) = \infty$ si $c = 0$
- Si $z = -\frac{d}{c}$, $m(z) = \infty$

La condition $ad - bc \neq 0$ nous permet d'éviter toute forme indéterminée.

On définit aussi $Möb^+$ comme l'ensemble des transformations de Möbius.

Proposition 3.2. *L'ensemble $Möb^+$ forme un groupe.*

Démonstration. Soit $m(z) = \frac{az+b}{cz+d}$ et $n(z) = \frac{\alpha z+\beta}{\gamma z+\delta}$ deux transformations de Möbius.

1. On a que $m \circ n(z)$ est une autre transformation de Möbius. En effet,

$$m \circ n(z) = \frac{a \frac{\alpha z+\beta}{\gamma z+\delta} + b}{c \frac{\alpha z+\beta}{\gamma z+\delta} + d} = \frac{a\alpha z + a\beta + b\gamma z + b\delta}{c\alpha z + c\beta + d\gamma z + d\delta} = \frac{(a\alpha + b\gamma)z + (a\beta + b\delta)}{(c\alpha + d\gamma)z + (c\beta + d\delta)}. \quad (1)$$

2. Les transformations de Möbius sont des bijections et leurs inverses sont eux-mêmes des transformations de Möbius. Soit $z' = m(z)$, on a

$$\begin{aligned} (cz + d)z' &= az + b \Rightarrow z(cz' - a) = b - dz' \\ \Rightarrow z &= \frac{b - dz'}{cz' - a} \\ \Rightarrow m^{-1}(z) &= \frac{b - dz}{cz - a}. \end{aligned}$$

De plus, la fonction identité est une transformation de Möbius et la composition de fonction est associative. Donc, les transformations de Möbius forment un groupe. \square

Remarque 3.3. Dans la preuve du théorème précédent, on voit un lien fort entre la multiplication de matrice 2×2 et la composition de transformations de Möbius, qui nous mène à la proposition suivante, qui permet de calculer plus efficacement la composition de deux transformations..

Proposition 3.4. *Il existe un morphisme trivial ι de $GL_2(\mathbb{C})$, les matrices 2×2 inversibles à coefficients complexes, vers le groupe $M\ddot{o}b^+$ des transformations de Möbius défini par :*

$$\iota : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \rightarrow \frac{az + b}{cz + d}.$$

Démonstration. Il suffit d'observer dans l'équation 1, qui calcule la composition de deux transformations, que lorsque l'on compose deux transformations, on obtient les mêmes coefficients que lorsque l'on multiplie deux matrices. \square

Note 3.5. Le morphisme ι n'est pas un isomorphisme. En effet, deux matrices peuvent représenter une seule transformation de Möbius. Par exemple, $2I_{2 \times 2} \neq I_{2 \times 2}$, mais ces deux matrices représentent la transformation identité.

La démonstration du théorème suivant est présentée dans [And05]. Ce théorème nous sera utile à quelques reprises.

Théorème 3.6. *Les transformations de Möbius sont des transformations conformes, c'est-à-dire qu'elles préservent les angles entre les tangentes à deux courbes à leur points d'intersections.*

Nous allons maintenant montrer que les transformations de Möbius envoient des cercles généralisés vers des cercles généralisés. Pour ce faire, nous allons prendre certaines transformations de Möbius assez simples et montrer qu'elles ont cette propriété. Puis, nous montrerons que toutes les transformations de Möbius s'expriment comme une composition de ces transformations de Möbius simples.

Lemme 3.7. *Les transformations de Möbius de la forme $m_0(z) = az + b$ où $a \neq 0$ pour $z \in \mathbb{C}$ et $m_0(\infty) = \infty$ envoient des cercles généralisés sur des cercles généralisés.*

Démonstration. Selon la proposition 2.5, un cercle généralisé c est défini par une équation de la forme $\alpha z\bar{z} + \beta z + \bar{\beta}\bar{z} + \gamma = 0$. Montrons que $\omega = az + b$ satisfait aussi à l'équation.

On a $z = \frac{1}{a}(\omega - b)$ en substituant dans l'équation du cercle, on obtient

$$\begin{aligned} 0 &= \alpha \frac{1}{a}(\omega - b) \frac{1}{\bar{a}}(\bar{\omega} - \bar{b}) + \frac{\beta}{a}(\omega - b) + \frac{\bar{\beta}}{\bar{a}}(\bar{\omega} - \bar{b}) + \gamma \\ &= \frac{\alpha}{|a|^2} \omega \bar{\omega} + \left(\frac{\beta}{a} - \frac{\alpha \bar{b}}{|a|^2} \right) \omega + \overline{\left(\frac{\beta}{a} - \frac{\alpha \bar{b}}{|a|^2} \right)} \bar{\omega} + (\gamma - 2 \operatorname{Re}(\frac{\beta}{a} b)) \end{aligned}$$

qui est de la forme de l'équation d'un cercle. \square

Remarque 3.8. Si c contient une droite euclidienne, son image par m_0 aussi, car $\frac{\alpha}{|a|^2} = 0 \Leftrightarrow \alpha = 0$. Ceci est cohérent avec le fait que $m_0(\infty) = \infty$, car un cercle généralisé contient une droite si, et seulement si il contient ∞ .

Lemme 3.9. *La fonction $m_1(z) = \frac{1}{z}$ pour $z \in \mathbb{C} \setminus \{0\}$, $m_1(0) = \infty$ et $m_1(\infty) = 0$ envoie des cercles sur des cercles.*

Démonstration. On fait comme le lemme précédent. Soit $\omega = \frac{1}{z} \Rightarrow z = \frac{1}{\omega}$, on a

$$\alpha \frac{1}{\omega} \frac{1}{\bar{\omega}} + \frac{\beta}{\omega} + \frac{\bar{\beta}}{\bar{\omega}} + \gamma = 0.$$

Si on multiplie par $\omega\bar{\omega}$, on a

$$\gamma\omega\bar{\omega} + \bar{\beta}\omega + \beta\bar{\omega} + \alpha = 0$$

qui est l'équation d'un cercle. \square

Théorème 3.10. *Les transformations de Möbius envoient des cercles sur des cercles dans $\hat{\mathbb{C}}$.*

Démonstration. Une transformation $m(z) = \frac{az+b}{cz+d}$ peut s'écrire comme :

- $m(z) = \frac{a}{d}z + \frac{b}{d}$ si $c = 0$;
- $m = n_0 \circ n_1 \circ n_2$ où $n_0(z) = -(ad-bc)z + \frac{a}{c}$, $n_1(z) = \frac{1}{z}$ et $n_2(z) = c^2z + cd$ si $c \neq 0$.

Dans les deux cas, on peut appliquer les deux lemmes précédents pour montrer que la transformation envoie des cercles sur des cercles. \square

On peut maintenant introduire un sous-groupe de $Möb^+$ préservant le demi-plan de Poincaré et y établir plusieurs propriétés qui nous seront utiles dans la section suivante.

Théorème 3.11. *L'ensemble*

$$Möb^+(\mathcal{H}) = \left\{ f(z) = \frac{az+b}{cz+d} \in Möb^+ \mid a, b, c, d \in \mathbb{R} \text{ et } ad-bc > 0 \right\}$$

est un sous-groupe de $Möb^+$ qui préserve \mathcal{H} .

Démonstration. Soit $m \in Möb^+(\mathcal{H})$ et $z \in \mathcal{H}$. On a,

$$\begin{aligned} \operatorname{Im}(m(z)) &= \operatorname{Im}\left(\frac{az+b}{cz+d}\right) \\ &= \operatorname{Im}\left(\frac{(az+b)(c\bar{z}+d)}{c^2+d^2}\right) \\ &= \frac{(ad-bc)\operatorname{Im}(z)}{c^2+d^2} > 0. \end{aligned}$$

On remarque puisque $a, b, c, d \in \mathbb{R}$, $z \notin \mathbb{R}$ et $z \neq \infty$, le numérateur et le dénominateur ne peuvent ni être nul, ni être l'infini. De plus, $ad-bc \neq 0 \Rightarrow c^2+d^2 \neq 0$. \square

Théorème 3.12. *L'ensemble $M\ddot{o}b^+(\mathcal{H})$ envoie des droites hyperboliques sur des droites hyperboliques.*

Démonstration. On sait que les transformations de Möbius envoient des cercles généralisés sur des cercles généralisés et préservent les angles entre deux courbes.

Comme les éléments de $M\ddot{o}b^+(\mathcal{H})$ préservent $\hat{\mathbb{R}}$, les cercles perpendiculaires à $\hat{\mathbb{R}}$ seront envoyés vers d'autres cercles perpendiculaires à $\hat{\mathbb{R}}$. Or, ceux-ci sont précisément des droites hyperboliques. \square

Théorème 3.13. *$M\ddot{o}b^+(\mathcal{H})$ agit transitivement sur l'ensemble des droites hyperboliques de \mathcal{H} .*

Démonstration. Soit d une droite hyperbolique. Selon le lemme 2.8, il suffit de trouver une transformation qui envoie cette droite sur l'axe imaginaire positif.

On pose x_0 et x_1 les points à l'infini de d .

Cas 1 : $x_0, x_1 \in \mathbb{R}$

On suppose sans perte de généralité que $x_1 > x_0$, on pose $m(z) = \frac{z-x_1}{z-x_0}$.

Cas 2 : $x_0 = \infty$ ou $x_1 = \infty$

On suppose sans perte de généralité que $x_0 = \infty$, on remarque que $x_1 \in \mathbb{R}$.

On pose $m(z) = z - x_1$.

Dans les deux cas on remarque que m envoie les points x_0 et x_1 sur ∞ et 0 , respectivement. Or, ce sont les points à l'infini de l'axe imaginaire positif, donc m envoie d sur l'axe imaginaire positif. Par conséquent, l'ensemble $M\ddot{o}b^+(\mathcal{H})$ agit transitivement sur l'ensemble des droites hyperboliques de \mathcal{H} . \square

Le précédent théorème peut être reformulé pour dire qu'étant donné deux droites hyperboliques, il existe au moins une transformation de Möbius qui envoie la première vers la deuxième.

4 Longueur hyperbolique

Cette section introduira la notion de longueur de chemin dans le demi-plan de Poincaré.

Définition 4.1. Soit $f : [a, b] \rightarrow \mathbb{R}^2$ continue dans $[a, b]$.

La fonction f est un chemin C^1 dans \mathbb{R}^2 si f est différentiable dans (a, b) avec sa dérivée continue.

La fonction f est un chemin C^1 par morceaux dans \mathbb{R}^2 s'il existe une partition de $[a, b]$ tel que f est un chemin C^1 si on le restreint à chaque élément de la partition.

Nous allons maintenant introduire une notation qui nous permettra de calculer des longueurs hyperboliques.

Notation 4.2. Soit $\mu : \mathbb{C} \rightarrow \mathbb{R}$ et $f : [a, b] \rightarrow \mathbb{R}^2$ un chemin C^1 . Si on voit f dans \mathbb{C} au lieu de \mathbb{R}^2 . C'est-à-dire $(x, y) \rightarrow x + yi$, on peut noter

$$\int_f \mu(z) |dz| = \int_a^b \mu(f(t)) |f'(t)| dt$$

où f' est la dérivée de f à valeur dans \mathbb{R}^2 .

On peut étendre cette notation à un chemin C^1 par morceaux en faisant la somme des intégrales sur les chemins C^1 d'une partition donnée par la définition du chemin C^1 par morceaux.

Le lemme suivant nous servira plus tard et peut être interprété intuitivement par le fait que si l'on parcourt un chemin au complet, on parcourra une distance supérieure ou égale à sa longueur (si l'on «reculait» lors du parcours, l'inégalité serait stricte).

Lemme 4.3. Soit $f : [a, b] \rightarrow \mathbb{C}$ un chemin par morceaux dans C^1 , $\rho : [a', b'] \rightarrow [a, b]$ une surjection dans C^1 et $\mu : \mathbb{C} \rightarrow \mathbb{R}^+$, alors

$$\int_f \mu(z) |dz| \leq \int_{f \circ \rho} \mu(z) |dz|.$$

De plus, il y a égalité si ρ est monotone.

Démonstration. On posera $s = \rho(t)$ et montrera seulement pour ρ , un chemin C^1 .

- Première partie, l'inégalité :

Puisque ρ est une surjection, il existe α et $\beta \in [a', b']$ tel que $\rho(\alpha) = a$ et $\rho(\beta) = b$. On supposera que $\beta < \alpha$, le cas $\beta > \alpha$ étant très similaire.

$$\begin{aligned} \int_{f \circ \rho} \mu(z) |dz| &= \int_{a'}^{b'} \mu((f \circ \rho)(t)) |(f \circ \rho)'(t)| dt \\ &\geq \int_{\beta}^{\alpha} \mu(f(\rho(t))) |f'(\rho(t))| |\rho'(t)| dt \quad \text{On restreint le domaine} \\ &\geq \left| \int_{\beta}^{\alpha} \mu(f(\rho(t))) |f'(\rho(t))| \rho'(t) dt \right| \quad \text{Inégalité des triangles} \\ &= \left| \int_b^a \mu(f(s)) |f'(s)| ds \right| = \int_f \mu(z) |dz|. \end{aligned}$$

- Deuxième partie, l'égalité si f est monotone :

Si $\rho'(t) \geq 0$ pour tout $t \in [a', b']$, alors $\rho(a') = a$, $\rho(b') = b$ et $|\rho'(t)| = \rho'(t)$.

$$\begin{aligned} \int_{f \circ \rho} \mu(z) |dz| &= \int_{a'}^{b'} \mu((f \circ \rho)(t)) |(f \circ \rho)'(t)| dt \\ &= \int_{a'}^{b'} \mu(f(\rho(t))) |f'(\rho(t))| |\rho'(t)| dt \\ &= \int_a^b \mu(f(s)) |f'(s)| ds = \int_f \mu(z) |dz|. \end{aligned}$$

Si $\rho'(t) \leq 0$ pour tout $t \in [a', b']$, alors $\rho(a') = b$, $\rho(b') = a$ et $|\rho'(t)| = -\rho'(t)$.

$$\begin{aligned} \int_{f \circ \rho} \mu(z) |dz| &= \int_{a'}^{b'} \mu(f(\rho(t))) |(f'(\rho(t)))| |\rho'(t)| dt \\ &= - \int_b^a \mu(f(s)) |(f'(s))| ds = \int_f \mu(z) |dz|. \end{aligned}$$

□

Remarque 4.4. Le lemme précédent nous permet de choisir comme on veut le domaine de définition d'un chemin C_1 par morceaux.

On a maintenant traité de tout ce qu'il faut pour introduire la notion de longueur hyperbolique.

Définition 4.5. Soit $\gamma > 0$, on définit la longueur hyperbolique d'un chemin C_1 par morceaux f contenu dans \mathcal{H} comme étant :

$$\text{longueur}_{\mathcal{H}}(f) = \int_f \frac{\gamma}{\text{Im}(z)} |dz|.$$

Pour la démonstration du prochain théorème, nous aurons besoin de la définition suivante qui étend la notion de dérivé et qui n'est pas utilisée ailleurs dans cet article.

Définition 4.6. Soit U un ouvert de \mathbb{C} et $f : U \rightarrow \mathbb{C}$, on définit la dérivée de f en z_0 comme suit : $f'(z_0) = \lim_{z \rightarrow z_0} \frac{f(z) - f(z_0)}{z - z_0}$ si cette limite existe. On dit que f est holomorphe en z_0 si cette limite existe.

Comme pour le cas réel, il y a des règles de dérivation. En particulier, nous utiliserons la règle de la composition dans la démonstration.

Théorème 4.7. Soit $f : [\alpha, \beta] \rightarrow \mathcal{H}$ un chemin C^1 par morceaux contenu dans \mathcal{H} et $\gamma > 0$. Sa longueur hyperbolique est invariante par rapport à $m(z) = \frac{az+b}{cz+d} \in \text{Möb}^+(\mathcal{H})$. C'est-à-dire :

$$\int_{m \circ f} \frac{\gamma}{\text{Im}(z)} |dz| = \int_f \frac{\gamma}{\text{Im}(z)} |dz|.$$

Démonstration. On montrera pour f un chemin C^1 seulement. On a que

$$m'(z) = \frac{ad - bc}{(cz + d)^2}$$

et

$$\begin{aligned} \text{Im}(m(z)) &= \text{Im} \left(\frac{az + b}{cz + d} \right) \\ &= \text{Im} \left(\frac{(az + b)(c\bar{z} + d)}{|cz + d|^2} \right) \\ &= \frac{(ad - bc) \text{Im}(z)}{|cz + d|^2}. \end{aligned}$$

D'où,

$$\begin{aligned}
\text{longueur}_{\mathcal{H}}(m \circ f) &= \int_{m \circ f} \frac{\gamma}{\text{Im}(z)} |dz| \\
&= \int_{\alpha}^{\beta} \frac{\gamma}{\text{Im}((m \circ f)(t))} |(m \circ f)'(t)| dt \\
&= \int_{\alpha}^{\beta} \frac{\gamma}{\text{Im}(m(f(t)))} |(m'(f(t)))| |f'(t)| dt \\
&= \int_{\alpha}^{\beta} \frac{\gamma |cf(t) + d|^2}{(ad - bc) \text{Im}(f(t))} \left| \frac{ad - bc}{(cf(t) + d)^2} \right| |f'(t)| dt \\
&= \int_{\alpha}^{\beta} \frac{\gamma}{\text{Im}(f(t))} |f'(t)| dt = \text{longueur}_{\mathcal{H}}(f).
\end{aligned}$$

□

Proposition 4.8. Soit $f : [a, b] \rightarrow \mathcal{H}$ un chemin par morceaux C^1 . Sa longueur hyperbolique est finie.

Démonstration. On montrera seulement pour f un chemin C^1 . Puisque $[a, b]$ est compacte, $f[a, b]$ est compacte et il existe $B > 0$ tel que

$$f[a, b] \subseteq \{z \in \mathcal{H} \mid \text{Im}(z) \geq B\}.$$

Puisque $|f'(t)|$ est continue et définie sur un intervalle fermé, par le théorème des bornes, il existe $d \in [a, b]$ tel que

$$|f'(t)| \leq |f'(d)| \quad \forall t \in [a, b].$$

D'où,

$$0 \leq \text{longueur}_{\mathcal{H}}(f) = \int_a^b \frac{\gamma}{\text{Im}(f(t))} |f'(t)| dt \leq \int_a^b \frac{\gamma}{B} |f'(d)| dt = \frac{\gamma |f'(d)|}{B} (b - a)$$

qui est finie.

□

5 Métrique de Poincaré

Cette section parlera de la métrique de Poincaré, qui permet de définir la distance hyperbolique entre deux points du demi-plan de Poincaré. En plus, de la notion de longueur d'un chemin dans le demi-plan de Poincaré. Pour cette section et l'autre suivante, on supposera que le paramètre γ de la longueur hyperbolique vaut 1.

Définition 5.1. Une métrique sur un ensemble E est une fonction $d : E \times E \rightarrow \mathbb{R}$ tel que

1. $d(x, y) \geq 0 \quad \forall x, y \in E$ et $d(x, y) = 0 \Leftrightarrow x = y$;

2. $d(x, y) = d(y, x) \quad \forall x, y \in E$;
3. $d(x, z) \leq d(x, y) + d(y, z) \quad \forall x, y, z \in E$.

Définition 5.2. On définit $\Gamma[x, y]$ l'ensemble des chemins C^1 par morceaux $f : [a, b] \rightarrow \mathcal{H}$ tel que $f(a) = x$ et $f(b) = y$.

Définition 5.3. On définit la fonction

$$d_{\mathbb{H}}(x, y) = \inf\{\text{longueur}_{\mathcal{H}}(f) \mid f \in \Gamma[x, y]\}.$$

On nommera cette fonction la métrique de Poincaré en vue d'un résultat à venir.

Proposition 5.4. La fonction $d_{\mathcal{H}}$ est invariante par rapport à $Möb^+(\mathcal{H})$, c'est-à-dire que si $m \in Möb^+(\mathcal{H})$ et $x, y \in \mathcal{H}$, $d_{\mathcal{H}}(x, y) = d_{\mathcal{H}}(m(x), m(y))$.

Démonstration. On remarque que $\{m \circ f \mid f \in \Gamma[x, y]\} \subseteq \Gamma[m(x), m(y)]$. On rappelle que la longueur hyperbolique est invariante par rapport à m . D'où

$$\begin{aligned} d_{\mathcal{H}}(m(x), m(y)) &= \inf\{\text{longueur}_{\mathcal{H}}(g) \mid g \in \Gamma[m(x), m(y)]\} \\ &\leq \inf\{\text{longueur}_{\mathcal{H}}(m \circ f) \mid f \in \Gamma[x, y]\} \\ &= \inf\{\text{longueur}_{\mathcal{H}}(f) \mid f \in \Gamma[x, y]\} = d_{\mathcal{H}}(x, y). \end{aligned}$$

Comme $m^{-1} \in Möb^+(\mathcal{H})$, on a $\{m^{-1} \circ g \mid g \in \Gamma[m(x), m(y)]\} \subseteq \Gamma[x, y]$ et effectuant les mêmes opérations, on obtient l'inégalité inverse. \square

Le théorème suivant nous permettra de calculer la distance hyperbolique entre deux points.

Théorème 5.5. On a l'égalité $d_{\mathcal{H}}(x, y) = \text{longueur}_{\mathcal{H}}(d)$ où d est une paramétrisation du segment de droite hyperbolique reliant x et y .

Démonstration. Puisque les transformations de Möbius agissent transitivement sur les droites hyperboliques et que $d_{\mathcal{H}}$ est invariant par rapport à ces transformations, on peut supposer que $x = ai$ et $y = bi$ où $a, b \in \mathbb{R}_+^*$ et, sans perte de généralité, que $a < b$.

Le chemin $f_0 : [a, b] \rightarrow \mathcal{H}$ défini par $f_0(t) = ti$ est une paramétrisation du segment de droite reliant x et y . On remarque que $\text{Im}(f_0(t)) = t$ et $|f'_0(t)| = 1$, donc

$$\text{longueur}_{\mathcal{H}}(f_0) = \int_a^b \frac{1}{t} dt = \ln\left(\frac{b}{a}\right).$$

Soit $f : [\alpha, \beta] \rightarrow \mathcal{H} \in \Gamma[x, y]$ défini par $f(t) = x(t) + y(t)i$, on veut que $\text{longueur}_{\mathcal{H}}(f_0) \leq \text{longueur}_{\mathcal{H}}(f)$. D'abord, on montre que l'on peut ignorer la partie réelle de f . Soit $g(t) = y(t)i$, on veut que la longueur de g soit plus courte que la longueur de f . On a $g(\alpha) = x$ et $g(\beta) = y$, donc $g \in \Gamma[x, y]$, mais

$$\begin{aligned} \text{longueur}_{\mathcal{H}}(g) &= \int_{\alpha}^{\beta} \frac{1}{y(t)} \sqrt{(y'(t))^2} dt \\ &\leq \int_{\alpha}^{\beta} \frac{1}{y(t)} \sqrt{(y'(t))^2 + (x'(t))^2} dt \\ &= \int_{\alpha}^{\beta} \frac{1}{\text{Im}(f(t))} |f'(t)| dt = \text{longueur}_{\mathcal{H}}(f). \end{aligned}$$

Il reste à montrer que $\text{longueur}_{\mathcal{H}}(f_0) \leq \text{longueur}_{\mathcal{H}}(g)$.

On a que $g([\alpha, \beta])$ est un segment de droite joignant $a'i$ et $b'i$ avec $a' \leq a < b \leq b'$. Soit $g_0 : [a', b'] \rightarrow \mathcal{H}$ défini par $g_0(t) = ti$. On a,

$$\text{longueur}_{\mathcal{H}}(f_0) = \ln \left(\frac{b}{a} \right) \leq \ln \left(\frac{b'}{a'} \right) = \text{longueur}_{\mathcal{H}}(g_0).$$

De plus, on remarque que $g = g_0 \circ (g_0^{-1} \circ g)$ où $(g_0^{-1} \circ g) : [\alpha, \beta] \rightarrow [a', b']$ est surjective. Donc, par le lemme 4.3, $\text{longueur}_{\mathcal{H}}(g_0) \leq \text{longueur}_{\mathcal{H}}(g)$, d'où

$$\text{longueur}_{\mathcal{H}}(f_0) \leq \text{longueur}_{\mathcal{H}}(g_0) \leq \text{longueur}_{\mathcal{H}}(g) \leq \text{longueur}_{\mathcal{H}}(f).$$

□

Il reste à vérifier que la fonction $d_{\mathcal{H}}$ est bien une métrique.

Théorème 5.6. *La métrique de Poincaré est une métrique.*

Démonstration. On vérifie les trois conditions.

1. Il est évident que $d_{\mathcal{H}}(x, y) \geq 0$ pour tout $x, y \in \mathcal{H}$. Soit $f : [a, b] \rightarrow \mathcal{H} \in \Gamma[x, y]$, on a $\text{longueur}_{\mathcal{H}}(f) = \int_a^b \frac{1}{\text{Im}(f(t))} |f'(t)| dt \geq 0$, car $\frac{1}{\text{Im}(f(t))}$ et $|f'(t)|$ sont positifs. Par conséquent, $d_{\mathcal{H}}(x, y) = \inf\{\text{longueur}_{\mathcal{H}}(f) \mid f \in \Gamma[x, y]\} \geq 0$.
2. On veut $d_{\mathcal{H}}(x, y) = d_{\mathcal{H}}(y, x)$. Il suffit de montrer que

$$\{\text{longueur}_{\mathcal{H}}(f) \mid f \in \Gamma[x, y]\} = \{\text{longueur}_{\mathcal{H}}(f) \mid f \in \Gamma[y, x]\}.$$

Ces deux ensembles auront alors le même infimum.

Soit $f : [a, b] \rightarrow \mathcal{H} \in \Gamma[x, y]$ et la surjection $\rho : [a, b] \rightarrow [a, b]$ définit par $\rho(t) = (a + b) - t$. Posons $f' = f \circ \rho$, on remarque que $f'(a) = y$ et $f'(b) = x$, donc $f' \in \Gamma[y, x]$. Or, puisque ρ est surjective et décroissante, selon le lemme 4.3, $\text{longueur}_{\mathcal{H}}(f) = \text{longueur}_{\mathcal{H}}(f')$ et, donc

$$\{\text{longueur}_{\mathcal{H}}(f) \mid f \in \Gamma[x, y]\} \subseteq \{\text{longueur}_{\mathcal{H}}(f) \mid f \in \Gamma[y, x]\}.$$

On peut montrer l'inclusion inverse par le même argument.

3. L'inégalité du triangle

Soit $x, y, z \in \mathcal{H}$ et $f_x : [a, b] \rightarrow \mathcal{H}$, $f_z : [b, c] \rightarrow \mathcal{H}$ des paramétrisations des segments de droites reliant x et y , et y et z , respectivement (on utilise ici la remarque 4.4). Soit $f : [a, c] \rightarrow \mathcal{H}$ la concaténation de f_x et f_z , on remarque que $f_x \in \Gamma[x, y]$, $f_z \in \Gamma[y, z]$ et $f \in \Gamma[x, z]$. De plus, selon la proposition 5.5, $d_{\mathcal{H}}(x, y) = \text{longueur}_{\mathcal{H}}(f_x)$ et $d_{\mathcal{H}}(y, z) = \text{longueur}_{\mathcal{H}}(f_z)$. D'où,

$$\begin{aligned} d_{\mathcal{H}}(x, z) &\leq \text{longueur}_{\mathcal{H}}(f) \\ &= \text{longueur}_{\mathcal{H}}(f_x) + \text{longueur}_{\mathcal{H}}(f_z) \\ &= d_{\mathcal{H}}(x, y) + d_{\mathcal{H}}(y, z). \end{aligned}$$

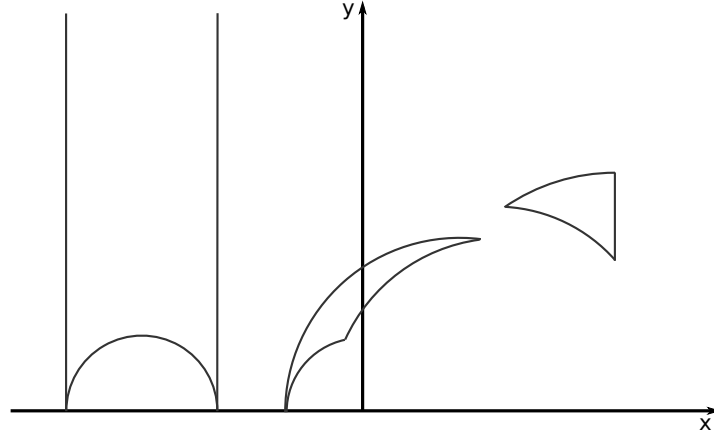


Figure 2 – Trois exemples de triangles hyperboliques. Le triangle de droite est compact

□

La définition suivante nous servira à démontrer le résultat final et donne un sens à la notion de distance entre deux ensembles.

Définition 5.7. Soit E un ensemble muni d'une métrique d . Soit aussi $X, Y \subset E$ non vide, et $e \in E$. On définit :

$$\begin{aligned} d(X, Y) &= \inf\{d(x, y) \mid x \in X, y \in Y\}; \\ d(e, X) &= \inf\{d(e, x) \mid x \in X\}. \end{aligned}$$

La distance entre X et Y , et e et X , respectivement.

6 Triangles hyperboliques et résultat final

Cette section étend la notion de triangle au demi-plan de Poincaré et introduit la trigonométrie dans la géométrie hyperbolique. Les triangles hyperboliques serviront à démontrer le résultat final de cet article. Lorsque l'on parlera d'angle entre deux courbes, nous parlerons de l'angle entre leurs tangentes.

Définition 6.1. Soit $A, B, C \in \overline{\mathcal{H}} = \mathcal{H} \cup \widehat{\mathbb{R}}$ tous différents. On définit le triangle hyperbolique $\triangle ABC$ la figure géométrique formée par les segments de droites, les demi-droites ou les droites hyperboliques reliant ces points, que l'on nomme côtés du triangle. On dit que $\triangle ABC$ est un triangle hyperbolique compact si $A, B, C \in \mathcal{H}$.

La figure 2 donne trois exemples de triangles hyperboliques. Le triangle de droite est le seul qui soit compact. Celui de gauche a comme extrémités trois points dans $\widehat{\mathbb{R}}$; de telles triangles sont dits idéals. Un triangle compact est donc formé de trois segments de droites joints à leurs extrémités de façon analogue

aux triangles euclidiens. Dans un triangle non compact, par contre, les segments de droites peuvent se prolonger et se rencontrer «à l'infini». Il est aussi possible d'avoir des triangles hyperboliques rectangles

Définition 6.2. Soit $\triangle ABC$ un triangle hyperbolique compact et α , β et γ ses angles intérieurs. On dit que $\triangle ABC$ est un triangle hyperbolique rectangle si α , β ou γ est un angle droit.

Proposition 6.3. Soit $\triangle ABC$ un triangle hyperbolique, au plus un de ses angles intérieurs est un angle droit.

Démonstration. Supposons que α et β sont deux angles droits du triangle $\triangle ABC$.

Puisque $Möb^+(\mathcal{H})$ agit transitivement sur les droites hyperboliques et que ses éléments préservent les angles, on peut supposer que le côté c entre α et β est sur l'axe imaginaire positif. Puisque α et β sont des angles droits, les deux autres côtés du triangle sont sur des demi-cercles distincts centrés en zéro. En particulier, ces côtés n'ont aucun point d'intersection, ce qui contredit la construction du $\triangle ABC$. \square

Définition 6.4. La proposition précédente nous permet de définir sans ambiguïté l'hypoténuse comme le côté opposé à l'angle droit et les cathètes comme les deux autres côtés.

On a, dans les triangles hyperboliques compacts, des relations qui sont très similaires aux relations dans les triangles euclidiens. Nous n'explorerons, par contre, que sommairement celles-ci.

Théorème 6.5. *Première loi des cosinus hyperbolique*

Soit $\triangle ABC$ un triangle hyperbolique compact avec α , β et γ comme angles intérieurs, et a , b et c , respectivement, les longueurs hyperboliques des côtés opposés à ces angles. On a la relation

$$\cosh(c) = \cosh(a) \cosh(b) - \sinh(a) \sinh(b) \cos(\gamma).$$

Contrairement au cas euclidien, il y a deux lois des cosinus. Cependant, nous n'aurons pas besoin de la seconde. La démonstration de la première loi des cosinus hyperbolique est trop longue pour l'utilisation que nous en ferons. Elle est présentée dans [And05] avec les autres relations dans les triangles et se fait dans un autre modèle de géométrie hyperbolique que le demi-plan de Poincaré.

Corollaire 6.6. *Théorème de Pythagore hyperbolique*

Soit $\triangle ABC$ un triangle rectangle hyperbolique d'hypoténuse de longueur c et de cathètes de longueur a et b . On a que l'hypoténuse est le côté le plus long du $\triangle ABC$ et la formule suivante :

$$\cosh(c) = \cosh(a) \cosh(b).$$

Démonstration. Pour la formule, on utilise la première loi des cosinus hyperbolique avec $\gamma = \frac{\pi}{2}$. On en déduit que l'hypoténuse est le côté le plus long, car $\cosh(x) \geq 1 \forall x \in \mathbb{R}^+$ et $\cosh(x)$ est croissante. \square

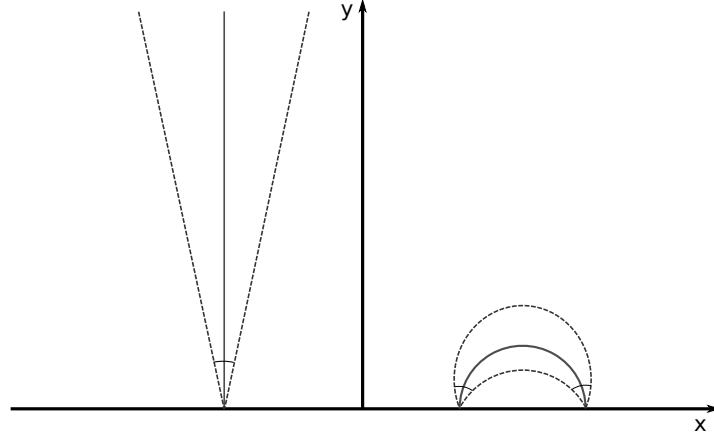


Figure 3 – Illustration du théorème 6.7

On a maintenant tous les résultats nécessaires pour démontrer le résultat final.

Théorème 6.7. *Soit \mathbf{D} une droite hyperbolique dans \mathcal{H} et $\delta \in \mathbb{R}_+^*$. Le lieu des points z tel que $d_{\mathcal{H}}(z, \mathbf{D}) = \delta$ forme deux arcs de cercle dans $\hat{\mathbb{C}}$ et les cercles contenant ces arcs intersectent les points à l'infini de \mathbf{D} avec le même angle.*

La figure 3 illustre le précédent théorème.

Démonstration. Puisque $Möb^+(\mathcal{H})$ agit transitivement sur les droites hyperboliques et que ses éléments envoient des cercles sur des cercles, on peut supposer que \mathbf{D} est l'axe imaginaire positif.

Soit $Z = \rho e^{i\theta} \in \mathcal{H} \setminus \mathbf{D}$, $P_0 = \rho i \in \mathbf{D}$ et $P \in \mathbf{D}$ tel que $P \neq P_0$. On remarque que $\triangle ZP_0P$ est un triangle rectangle et que son hypoténuse est le segment de droite hyperbolique reliant Z et P . Par conséquent, le segment de droite \mathbf{S} reliant Z et P_0 est plus court et, puisque l'on a pris un point P quelconque sur \mathbf{D} , $d_{\mathcal{H}}(Z, \mathbf{D}) = \text{longueur}_{\mathcal{H}}(f)$ où f est une paramétrisation de \mathbf{S} . La figure 4 donne une illustration du raisonnement.

On remarque que \mathbf{S} est sur un cercle centré en zéro.

Si $\theta < \frac{\pi}{2}$, une paramétrisation de \mathbf{S} est $f : [\theta, \frac{\pi}{2}] \rightarrow S$ défini par $f(t) = \rho e^{it}$. On remarque que $|f'(t)| = |\rho i e^{it}| = \rho$ et que $\text{Im}(f(t)) = \rho \sin t$. On veut

$$\begin{aligned} \delta &= \text{longueur}_{\mathcal{H}}(f) \\ &= \int_{\theta}^{\frac{\pi}{2}} \frac{1}{\rho \sin t} \rho dt \\ &= \ln(\text{cosec}(\theta) - \cotg(\theta)). \end{aligned}$$

Si $\theta > \frac{\pi}{2}$, une paramétrisation de \mathbf{S} est $f : [\frac{\pi}{2}, \theta] \rightarrow S$ défini par $f(t) = \rho e^{it}$ et on obtient

$$\delta = -\ln(\text{cosec}(\theta) - \cotg(\theta)).$$

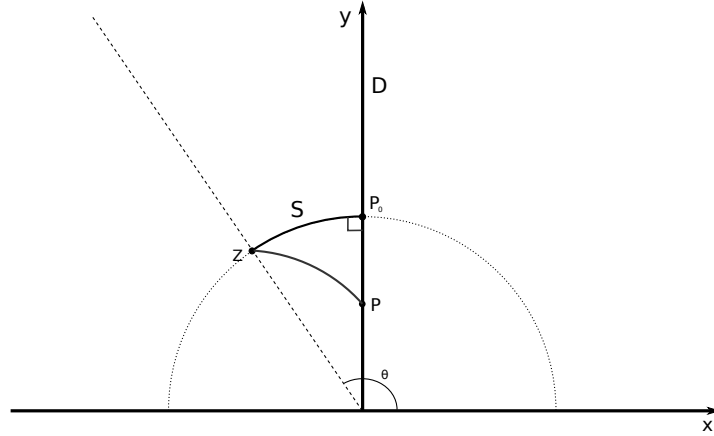


Figure 4 – Illustration de la démonstration du théorème 6.7

Dans les deux cas on a que δ est une fonction de θ seulement et que $\delta(\theta)$ est une bijection de $(0, \frac{\pi}{2})$ et $(\frac{\pi}{2}, \pi)$, respectivement, vers \mathbb{R}_+^* . Le lieu des points z tel que $d_{\mathcal{H}}(z, D) = \delta$ forme donc deux demi-droites avec comme extrémité zéro. Donc, les cercles les contenant intersectent les points à l'infini de \mathbf{D} , soit 0 et ∞ . En remplaçant θ par $\pi - \theta$, on peut montrer qu'ils forment le même angle par rapport à \mathbf{D} . Puisque $Möb^+(\mathcal{H})$ envoie des cercles généralisés vers cercles généralisés, ces lieux de points sont toujours des arcs de cercle.

On remarque que puisque \mathbb{R}_+^* préservent les angles, il y aura toujours le même angle θ entre la droite et les deux arcs de cercle. De plus, si la droite est sur un cercle euclidien et donc les arcs aussi, les angles aux deux points à l'infini de la droite entre la droite et chaque arc doivent géométriquement être égaux. \square

Ce résultat illustre bien le fait que la géométrie hyperbolique est fort différente de la géométrie euclidienne. En effet, dans le cas euclidien ces lieux de points correspondent aux droites euclidiennes. Dans le cas hyperbolique, au contraire, ce ne sont pas des droites hyperboliques, car les arcs de cercle ne forment pas un angle droit avec l'axe réel. La géométrie hyperbolique a donc certaines propriétés qui échappent à notre intuition euclidien. Nous n'avons vu que certaines d'entre elles. Elle est donc un bel exemple de la richesse que peut avoir la géométrie.

Références

[And05] James ANDERSON : *Hyperbolic Geometry*. Springer, 2005.

KEVIN THOUIN
DÉPARTEMENT DE MATHÉMATIQUES, UNIVERSITÉ DE SHERBROOKE
Courriel: Kevin.Thouin@USherbrooke.ca

La structure d’algèbre amassée des grassmaniennes $\text{Gr}(k, n)$

Souheila Hassoun, Denis Langford, Frédéric Langlois

Reçu le 2017-07-06 et accepté le 2017-11-23

RÉSUMÉ La grassmannienne $\text{Gr}(k, n)$, qui est l’un des objets étudiés dans le domaine de la géométrie algébrique, est l’ensemble de tous les sous-espaces de dimension k d’un espace vectoriel de dimension n . Il se trouve que, pour chaque grassmannienne, l’anneau de coordonnées d’une grassmannienne est isomorphe à une algèbre amassée, un objet très populaire dans la théorie de représentation. Dans cet article, pour $\text{Gr}(3, 6)$, nous avons calculé toutes les algèbres inclinées et inclinées amassées. Nous avons ensuite vérifié la validité d’un magnifique théorème nous venant de la théorie de représentation qui fait un lien entre ces algèbres et l’algèbre amassée associée à cette grassmannienne.

1 Introduction

Cet article résume le travail de notre groupe durant l’hiver 2017 sous la direction du professeur Thomas Brüstle, qui portait sur l’étude de la structure d’algèbre amassée des grassmaniennes.

Dans la deuxième section de cet article, nous allons décrire ce qu’est un carquois, un amas, une graine, une mutation d’une graine dans une certaine direction, puis, en utilisant ces notions, nous allons définir l’algèbre amassée générée par une graine S . Dans cette même section, une partie sera consacrée à l’algèbre amassée de type A_2 pour illustrer la théorie que nous avons déjà introduite et quelques propriétés intéressantes des algèbres amassées comme le phénomène de Laurent et la relation entre le type d’une algèbre amassée et la forme de son carquois initial. Nous y introduirons aussi les carquois de Dynkin de types A_n , D_n , E_6 , E_7 , et E_8 .

Dans la troisième section, nous définirons et étudierons les grassmaniennes. Nous énoncerons le théorème de Julius Plücker concernant l’anneau des coordonnées de la grassmannienne $\text{Gr}(2, 4)$ et nous définirons les variables de Plücker d’une grassmannienne $\text{Gr}(k, n)$ en général. En utilisant ces variables, nous énoncerons un théorème dû à Fomin et Zelevinsky concernant le cas $k = 2$ et à Scott pour le cas $k \geq 3$. Ce théorème indique que l’anneau des coordonnées d’une grassmannienne admet la structure d’une algèbre amassée. La grassmannienne $\text{Gr}(2, 4)$ sera amenée en exemple et nous consacrerons une sous-section à la description

Nous remercions grandement Thomas Brüstle d’avoir supervisé notre groupe de travail.

explicite de la grassmanienne $\text{Gr}(3,6)$ ainsi qu'à ses propriétés afin de vérifier les propriétés discutées dans la section précédente.

Dans la dernière section, nous parlerons des algèbres inclinées et inclinées amassées. Après avoir défini ces dernières, nous définirons la dimension globale d'une algèbre et caractériserons la dimension globale de ces algèbres. Nous donnerons l'exemple du carquois de type D_4 et nous observerons une bijection entre les carquois des algèbres inclinées amassées et les carquois des graines de la grassmanienne $\text{Gr}(3,6)$. Ces dernières sont calculées dans la Section 3. Cette bijection vérifie les résultats de Buan, Marsh, Reineke, Reiten et Todorov [BMR⁺06].

2 Algèbre amassée

La structure de l'anneau des coordonnées d'une grassmanienne est celle d'une algèbre amassée. Pour cette raison, nous commencerons par définir ce qu'est une algèbre amassée et nous donnerons un exemple de construction d'une telle algèbre.

Notons que les algèbres amassées ont été introduites et développées par Sergey Fomin et Andrei Zelevinsky [FZ02].

2.1 Définitions

Afin de comprendre la procédure de construction d'une algèbre amassée, quelques concepts et définitions sont nécessaires.

Les définitions qui suivent utilisent le concept de carquois. Un *carquois* est simplement un graphe orienté composé d'un ensemble de sommets associé à un ensemble de flèches qui relient un sommet à un autre sommet ou à lui-même. En d'autres termes, c'est un graphe dont les arêtes sont des flèches.

Exemple 2.1. Voici une représentation classique de carquois :

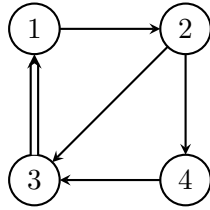


Figure 1 – Exemple de carquois

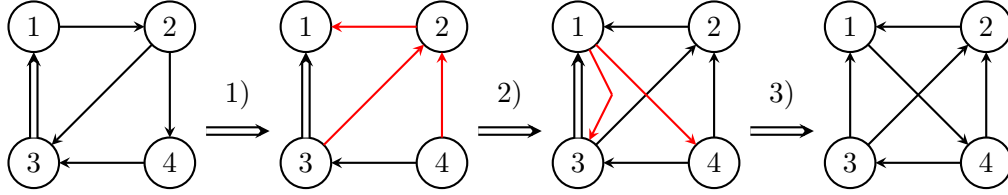
Ce carquois contient quatre sommets étiquetés 1, 2, 3 et 4 ; quatre flèches simples $1 \rightarrow 2$, $2 \rightarrow 3$, $2 \rightarrow 4$ et $4 \rightarrow 3$; et une flèche double $3 \Rightarrow 1$. Une flèche double est simplement un ensemble de deux flèches simples.

Note 2.2. Pour les algèbres amassées, on travaillera uniquement avec des carquois ne contenant aucune boucle et aucun 2-cycle.

Définition 2.3. Étant donné un carquois Q formé de n sommets, la mutation par le sommet k ($1 \leq k \leq n$) transforme le carquois Q en un nouveau carquois $Q' = m_k(Q)$ selon la liste de transformations suivantes :

1. Inverser toute flèche qui débute ou qui termine au sommet k ;
2. Pour tout chemin de la forme $i \longrightarrow k \longrightarrow j$, ajouter une nouvelle flèche $j \longrightarrow i$;
3. Retirer tous les 2-cycles.

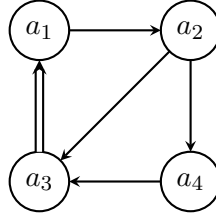
Exemple 2.4. Disons que Q est le carquois de l'exemple précédent (Figure 1). On veut $Q' = m_2(Q)$ (mutation du carquois Q par le sommet 2).



Définition 2.5. Soit l'ensemble de variables $\{a_1, a_2, \dots, a_n\}$ transcendent sur \mathbb{Q} . Alors la famille (a_1, a_2, \dots, a_n) est appelée un *amas* et a_1, a_2, \dots, a_n sont appelées des *variables amassées*.

Définition 2.6. Étant donné un amas x et un carquois Q , la paire $S := (x, Q)$ est appelée une *graine*.

Exemple 2.7. Disons que Q est le carquois de la figure 1 et posons l'amas $x := \{a_1, a_2, a_3, a_4\}$. Alors, la graine $S := (x, Q)$ serait la suivante :



Définition 2.8. Soit un amas $x := \{a_1, a_2, \dots, a_n\}$, un carquois Q et une graine $S := (x, Q)$. La *mutation de la graine S de direction k* est l'application $\mu_k(S) : S \longrightarrow S'$ où :

1. $S' := (x', Q')$ est une nouvelle graine formée d'un nouvel amas x' et d'un nouveau carquois Q' ;
2. $x' := (a'_1, a'_2, \dots, a'_n)$ où
si \mathcal{I} est l'ensemble des sommets $i \in Q$ tels qu'il existe une flèche $i \longrightarrow k$ et
si \mathcal{O} est l'ensemble des sommets $j \in Q$ tels qu'il existe une flèche $k \longrightarrow j$
alors,

$$a_t = \begin{cases} \frac{1}{a_t} [\prod_{i \in \mathcal{I}} a_i + \prod_{j \in \mathcal{O}} a_j], & \text{si } t = k \\ a_k, & \text{sinon} \end{cases}$$

3. $Q' = m_k(Q)$

Plusieurs exemples de cette opération sont disponibles dans la Section 2.2.

Remarque 2.9. Cette mutation est involutive, c'est-à-dire que si l'on mute deux fois dans la même direction, on revient à la graine initiale. Plus formellement : $m_k(m_k(Q)) = Q$.

Définition 2.10. Une variable amassée a est dite gelée si la mutation dans la direction de cette variable n'est pas permise. De même, elle est non-gelée si on peut muter dans sa direction.

Notation 2.11. Dans tous les carquois de cet article représentant des graines, les variables amassées seront indiquées par des carrés et les non-gelées par des cercles.

Définition 2.12. Soient un amas $x := (a_1, a_2, \dots, a_n)$, un carquois Q , et une graine $S := (x, Q)$.

On définit \mathcal{A} comme l'ensemble de toutes les variables d'amas obtenues par un nombre arbitraire de mutations sur S .

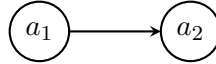
$$\mathcal{A} := \{a \in \mu_{i_1} \mu_{i_2} \cdots \mu_{i_r} S \mid 1 \leq i_1, i_2, \dots, i_r \leq n\}$$

On définit $\mathcal{C}(S)$, l'*algèbre amassée générée par la graine S* , comme étant le résultat de l'extension algébrique de \mathbb{Z} générée par \mathcal{A} .

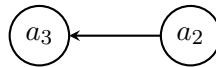
Les éléments de \mathcal{A} sont les *coordonnées* de $\mathcal{C}(S)$.

2.2 Construction d'une algèbre amassée de type A_2

Tel qu'indiqué dans la section précédente, les algèbres amassées sont construites par une série arbitraire de mutations sur une graine initiale. L'algèbre amassée de type A_2 s'appelle ainsi étant donnée qu'elle est construite à partir d'une graine initiale dont le carquois associé est un graphe de Dynkin de type A_2 . On choisit donc notre graine initiale $S_0 := (x_0, Q_0)$ avec l'amas initial $x := \{a_1, a_2\}$ et le graphe initial Q suivant :



On n'a maintenant que deux choix de mutations possibles. Sans perte de généralité, pour le présent article, nous avons choisi une mutation sur le sommet 1. À la suite de la mutation, on obtient une nouvelle graine $S_1 = (x_1, Q_1) = \mu_1(S_0)$ ayant la forme suivante :

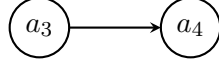


Où $x_1 = (a_3, a_2)$.

On calcule a_3 à l'aide de la formule fournie dans la Définition 2.8 :

$$a_3 = \frac{1}{a_1}(1 + a_2) = \frac{1 + a_2}{a_1}$$

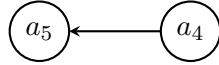
On a que $\mu_k(\mu_k(S)) = S$. Donc, pour la suite, nous ne pouvons qu'alterner la sélection du sommet pour les mutations. On trouvera donc une nouvelle graine $S_2 := (x_2, Q_0) = \mu_2(S_1)$:



Où $x_2 = (a_3, a_4)$, avec

$$a_4 = \frac{1}{a_2}(1 + a_3) = \frac{1}{a_2}\left(1 + \frac{1 + a_2}{a_1}\right) = \frac{1 + a_1 + a_2}{a_1 a_2}$$

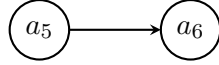
Dans le même ordre d'idées, on trouve la graine suivante $S_3 := (x_3, Q_1) = \mu_1(S_3)$:



Où $x_3 = (a_5, a_4)$, avec

$$a_5 = \frac{1}{a_3}(1 + a_4) = \frac{a_1}{1 + a_2}\left(1 + \frac{1 + a_1 + a_2}{a_1 a_2}\right) = \frac{1 + a_1}{a_2}$$

Ensuite, on trouve $S_4 := (x_4, Q_0) = \mu_2(S_3)$:



Où $x_4 = (a_5, a_6)$, avec

$$a_6 = \frac{1}{a_4}(1 + a_5) = \frac{a_1 a_2}{1 + a_1 + a_2}\left(1 + \frac{1 + a_1}{a_2}\right) = a_1$$

Enfin, on trouve $S_5 := (x_5, Q_1) = \mu_1(S_4)$:

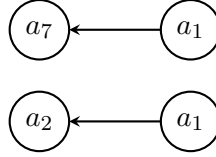
Où $x_5 = (a_7, a_1)$, avec

$$a_7 = \frac{1}{a_5}(1 + a_1) = \frac{a_2}{1 + a_1}(1 + a_1) = a_2$$

On peut maintenant noter que notre dernière graine S_5 a l'allure suivante :

On voit clairement qu'il existe un isomorphisme entre la graine S_5 et la graine S_0 . De plus, si l'on décidait de continuer notre manège de mutations plus longtemps, on retrouverait exactement quatre graines isomorphes à l'une des graines déjà obtenues pour ensuite revenir à la graine S_0 . Ainsi, on a trouvé toutes les graines existantes à isomorphisme près.

On a, de plus, trouvé toutes les variables d'amas possibles dans la situation actuelle. Ces dernières sont :



$$\left\{ a_1, a_2, \frac{1+a_1}{a_2}, \frac{1+a_2}{a_1}, \frac{1+a_1+a_2}{a_1 a_2} \right\} := \mathcal{A}$$

On obtient finalement l'algèbre amassée engendrée par la graine S_0 , $\mathcal{C}(S_0)$, en calculant $\mathbb{Z}(\mathcal{A})$, l'extension algébrique de l'ensemble \mathcal{A} .

2.3 Quelques propriétés intéressantes des algèbres amassées

Définition 2.13. Si \mathbb{K} est un corps et $k_i \in \mathbb{K}$. Alors, $\mathbb{L}[x] \in \mathbb{K}[x, x^{-1}]$ est un *polynôme de Laurent* s'il est un polynôme de la forme :

$$\mathbb{L}[x] = \left[\sum_{i \in \mathbb{Z}} k_i x^i \right]$$

Le théorème suivant présente un très beau résultat qui donne une expression algébrique bien particulière aux variables amassées.

Théorème 2.14. *Pour toute algèbre amassée, toute variable amassée obtenue par la mutation d'une graine initiale peut être représentée par un polynôme de Laurent ayant pour variables les variables amassées contenues dans l'amas initial. Toute variable amassée a est un polynôme de Laurent ayant pour variables $\{a_1, \dots, a_n, a_1^{-1}, \dots, a_n^{-1}\}$. On peut alors s'écrire :*

$$a \in \mathbb{L}[a_1, \dots, a_n, a_1^{-1}, \dots, a_n^{-1}]$$

Ce qui permet d'écrire que

$$a = \frac{\mathcal{P}(a_1, \dots, a_n)}{a_1^{p_1} \dots a_n^{p_n}}$$

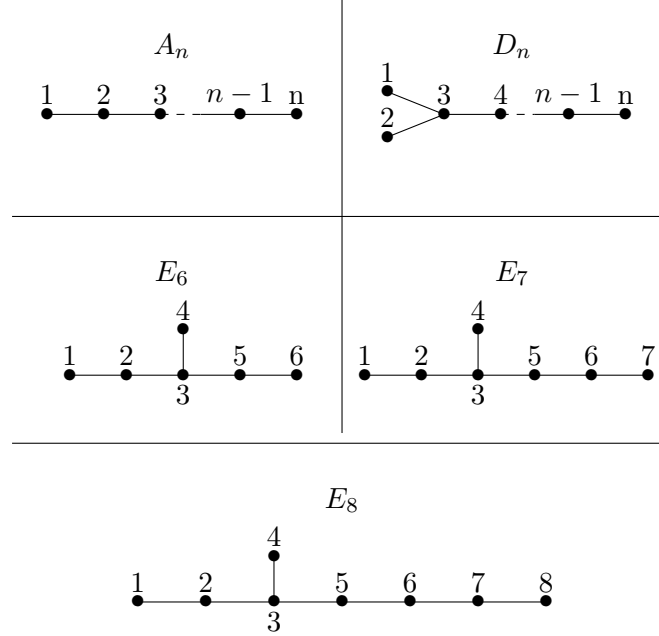
Ce dernier théorème, prouvé dans [FZ02], présente une belle expression pour toutes les variables amassées. Elles sont toutefois encore plus raffinées. Le théorème suivant en témoigne.

Théorème 2.15. *Pour toute algèbre amassée, les coefficients du polynôme au numérateur de $a = \frac{\mathcal{P}(a_1, \dots, a_n)}{a_1^{p_1} \dots a_n^{p_n}}$, une variable amassée, sont tous des entiers positifs.*

Définition 2.16. Une algèbre amassée est dite de type fini si \mathcal{A} , l'ensemble de toutes les variables amassées qu'elle engendre par la mutation, est lui-même fini.

Théorème 2.17. *Une algèbre amassée de carquois Q est de type fini si et seulement si Q peut être obtenu par mutation d'un carquois de type Dynkin.*

Les carquois de type Dynkin sont originaires de la théorie de Lie. Ils sont classifiés de la manière suivante :



Introduits par Eugene Dynkin, ces graphes sont classés selon des types bien précis. Les graphes présentés ci-dessus sont du type non-gradué. Il existe des graphes du type gradué qui sont étudiés eux aussi dans le contexte des algèbres amassées. Par exemple, les créateurs des algèbres amassées Sergey Fomin et Andrei Zelevinsky les ont étudiées dans [FZ02]. Nous nous limiterons cependant aux graphes non-gradués puisque c'est d'eux dont on a besoin pour explorer les grassmaniennes qui seront définies dans la prochaine section.

3 Grassmaniennes

3.1 Définitions

Définition 3.1. Une grassmannienne $\text{Gr}(k, n)$ est l'ensemble de tous les sous-espaces de \mathbb{C}^n qui sont de dimensions k .

La grassmannienne $\text{Gr}(k, n)$ est étudiée dans le domaine de la géométrie algébrique. Un théorème classique de Julius Plücker décrit son anneau de coordonnées comme étant le quotient d'un anneau de polynômes par certaines relations qui portent maintenant son nom.

Théorème 3.2 (Plücker). *Pour $(k,n) = (2,4)$,*

$$\mathbb{C}[\text{Gr}(k,n)] = \frac{\mathbb{C}[\chi_I \mid I \subset \{1,2,\dots,n\}, |I| = k]}{\text{relations de Plücker}}$$

où les relations de Plücker sont données par :

$$\chi_{1,3}\chi_{2,4} = \chi_{1,4}\chi_{2,3} + \chi_{1,2}\chi_{3,4}$$

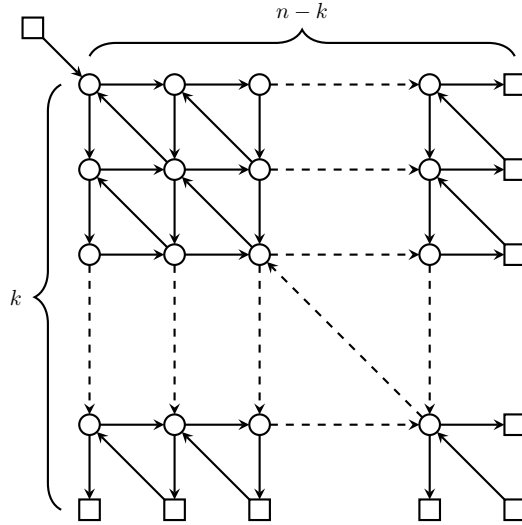
Définition 3.3. On nomme *variables de Plücker d'une grassmannienne $\text{Gr}(k,n)$* les variables contenues dans l'ensemble suivant :

$$\{\chi_I \mid I \subset \{1,2,\dots,n\}, |I| = k\}$$

Théorème 3.4 (Fomin, Zelevinsky pour $k=2$ /Scott pour $k \geq 3$). *Soient la grassmannienne $\text{Gr}(k,n)$ et son anneau de coordonnées $\mathbb{C}[\text{Gr}(k,n)]$.*

Soient X un ensemble de $(k-1)(n-k-1)$ variables non-gelées et F un ensemble de n variables gelées.

Soit Q le carquois composé des ensembles X et F et disposé de la manière suivante :



Si \mathcal{A}_Q est l'algèbre amassée obtenue par mutations sur Q , alors $\mathbb{C}[\text{Gr}(k,n)]$ est isomorphe \mathcal{A}_Q .

Ensuite, pour le cas $k = 2$, il existe une bijection entre les $\chi_{i,j}$ dans χ_I et les variables amassées.

Lorsque $k \geq 3$, \mathcal{A}_Q contient toujours plus de variables amassées qu'il y a de variables de Plücker.

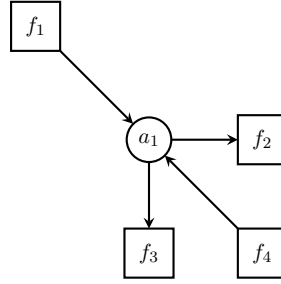
Enfin, soit Q' , le sous-carquois de Q n'incluant que les variables de l'ensemble X et les flèches les reliant. Tous les cas finis des algèbres amassées \mathcal{A}_Q sont connus et leur sous-carquois Q' est obtenu par mutations d'un graphe de Dynkin.

Le tableau suivant présente les cas finis (tous les autres cas sont infinis) :

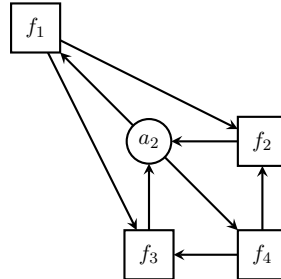
<i>Grassmannienne</i>	<i>Type Dynkin</i>
$\text{Gr}(2,n)$	A_{n-3}
$\text{Gr}(3,6)$	D_4
$\text{Gr}(3,7)$	E_6
$\text{Gr}(3,8)$	E_8

3.1.1 Court exemple : La grassmannienne $\text{Gr}(2,4)$

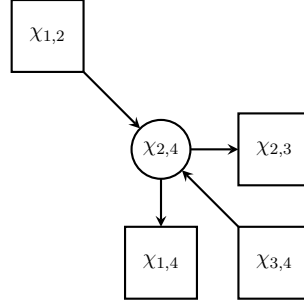
En général, les grassmanniennes sont des objets très complexes. Cet exemple de grassmannienne est heureusement assez simple. Prenons $\text{Gr}(2,4)$. Nous avons donc $\text{Gr}(2,4) = \{\mathcal{V} \subseteq \mathbb{C}^4 \mid \dim \mathcal{V} = 2\}$. Considérons l'algèbre amassée isomorphe à l'anneau de coordonnées de $\text{Gr}(2,4)$. Par le théorème 3.4, nous aurons la graine initiale qui sera de la forme suivante :



On note qu'il y a une seule variable non-gelée. On ne peut donc faire la mutation que sur cette variable. Par cette mutation, on obtient la nouvelle graine qui est de la forme suivante :



Comme la mutation est involutive, ce sont les deux seules graines de cette algèbre amassée. Nous pouvons associer à chaque sommet de la graine initiale une variable de Plücker comme suit :



On peut alors observer que, dans ce cas, la formule de mutation de a_1 donne $a_1 a_2 = f_1 f_4 + f_2 f_3$. Ceci correspond à la relation de Plücker : $\chi_{1,3} \chi_{2,4} = \chi_{1,4} \chi_{2,3} + \chi_{1,2} \chi_{3,4}$. On peut en déduire qu'il y a une bijection entre les variables de Plücker et les variables amassées pour $\text{Gr}(2,4)$ comme l'indique le théorème 3.4. Cette bijection est donnée par le tableau suivant :

<i>Variable amassée</i>	<i>Plücker associée</i>
a_1	$\chi_{2,4}$
a_2	$\chi_{1,3}$
f_1	$\chi_{1,2}$
f_2	$\chi_{2,3}$
f_3	$\chi_{1,4}$
f_4	$\chi_{3,4}$

L'exemple suivant est beaucoup plus complexe. Il nous permettra d'observer la validité du théorème 3.4. En effet, dans ce cas, il y a plus de variables amassées que de variables de Plücker.

3.1.2 La grassmannienne $\text{Gr}(3,6)$

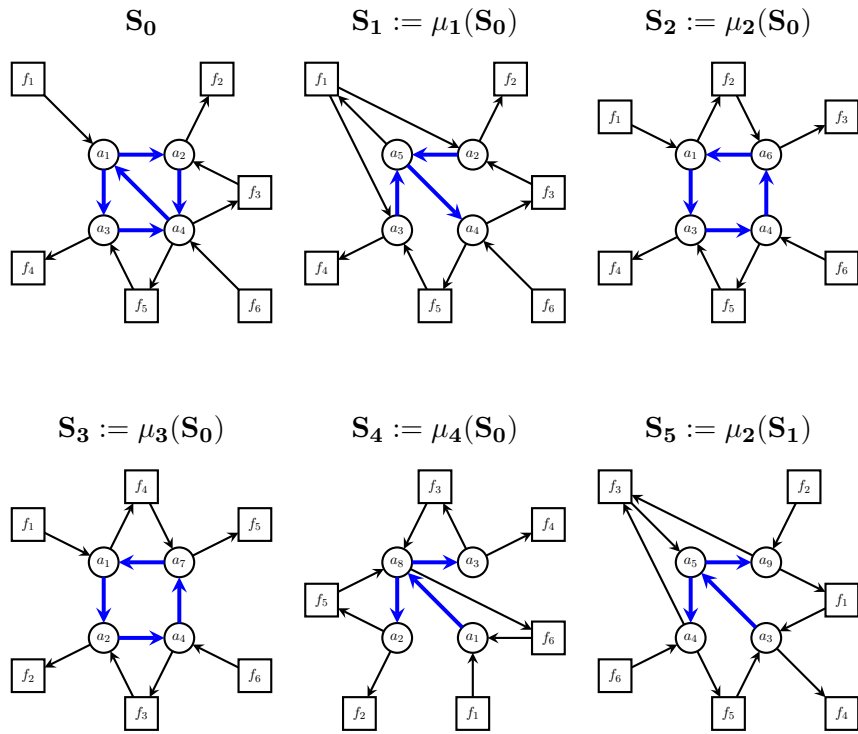
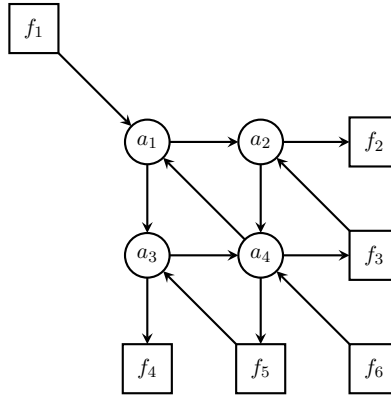
Pour construire la grassmannienne $\text{Gr}(3,6)$, nous devons établir une graine initiale de la forme indiquée par le théorème de Fomin, Zelevinsky et Scott (Théorème 3.4) :

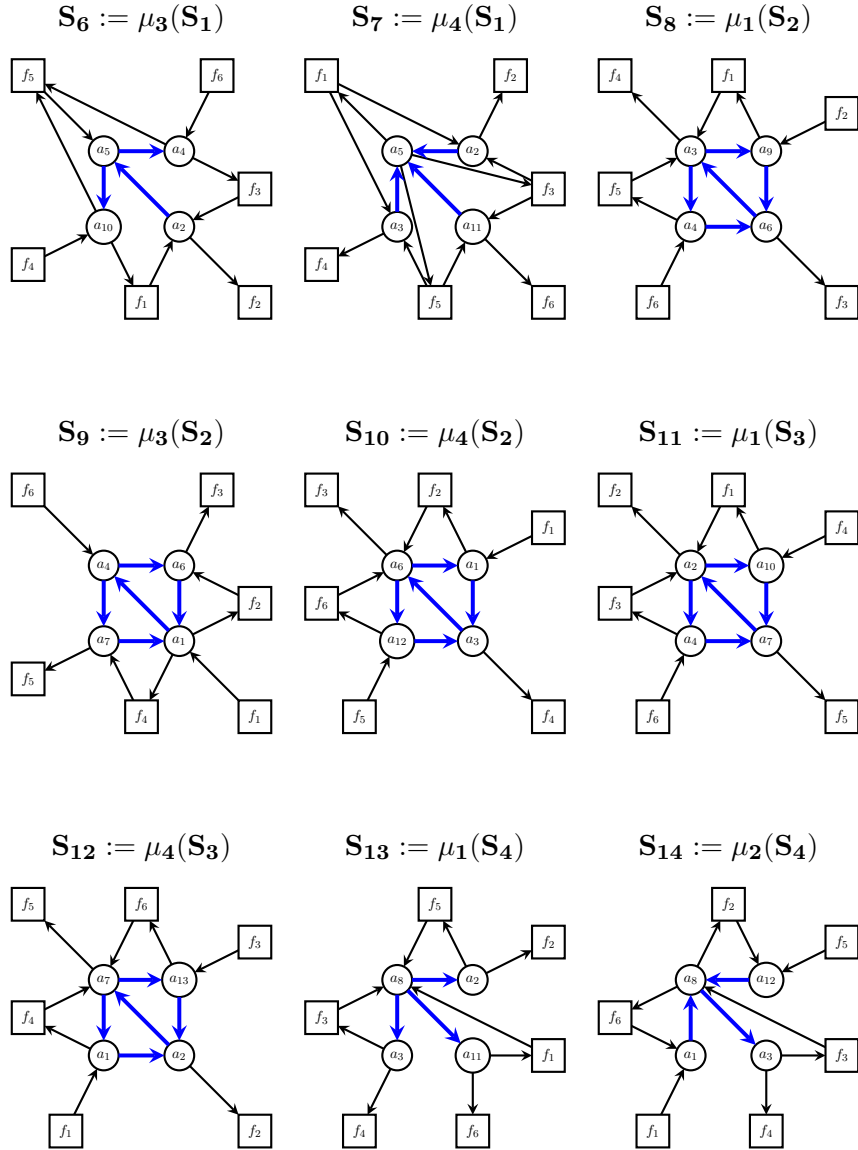
Soit un amas $\{a_1, a_2, a_3, a_4, f_1, f_2, f_3, f_4, f_5, f_6\} := x_0$.

Soit une graine $S_0 := (Q_0, x_0)$ ayant la forme suivante :

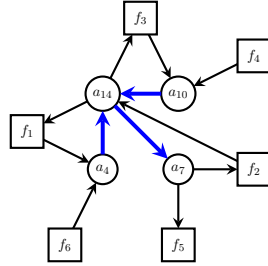
Alors, si les sommets carrés qui sont associés aux variables f_i sont gelés, le théorème 3.4 indique que l'on peut obtenir l'anneau des coordonnées de la grassmannienne $\text{Gr}(3,6)$ en appliquant toutes les combinaisons de mutations possibles sur S_0 .

Voici une liste complète de toutes les graines obtenues de cette façon (à isomorphisme de graphe près) :

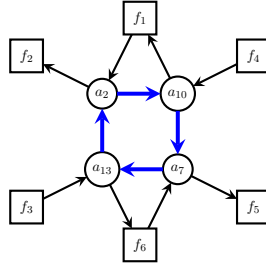




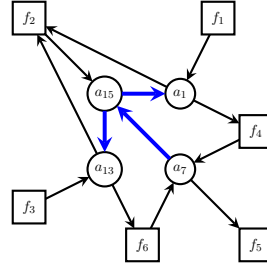
$$S_{24} := \mu_2(S_{11})$$



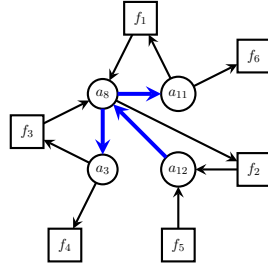
$$S_{25} := \mu_4(S_{11})$$



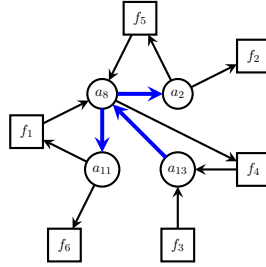
$$S_{26} := \mu_2(S_{12})$$



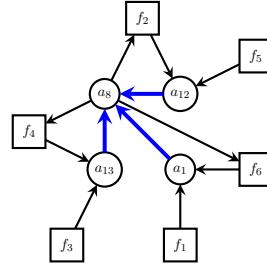
$$S_{27} := \mu_2(S_{13})$$



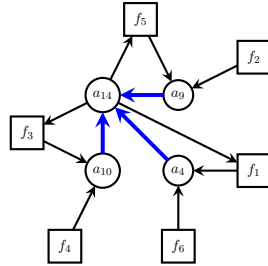
$$S_{28} := \mu_3(S_{13})$$



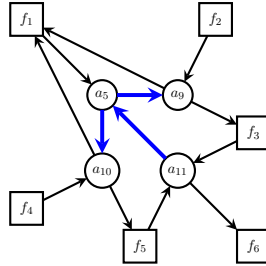
$$S_{29} := \mu_3(S_{14})$$



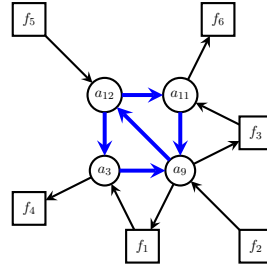
$$S_{30} := \mu_1(S_{16})$$

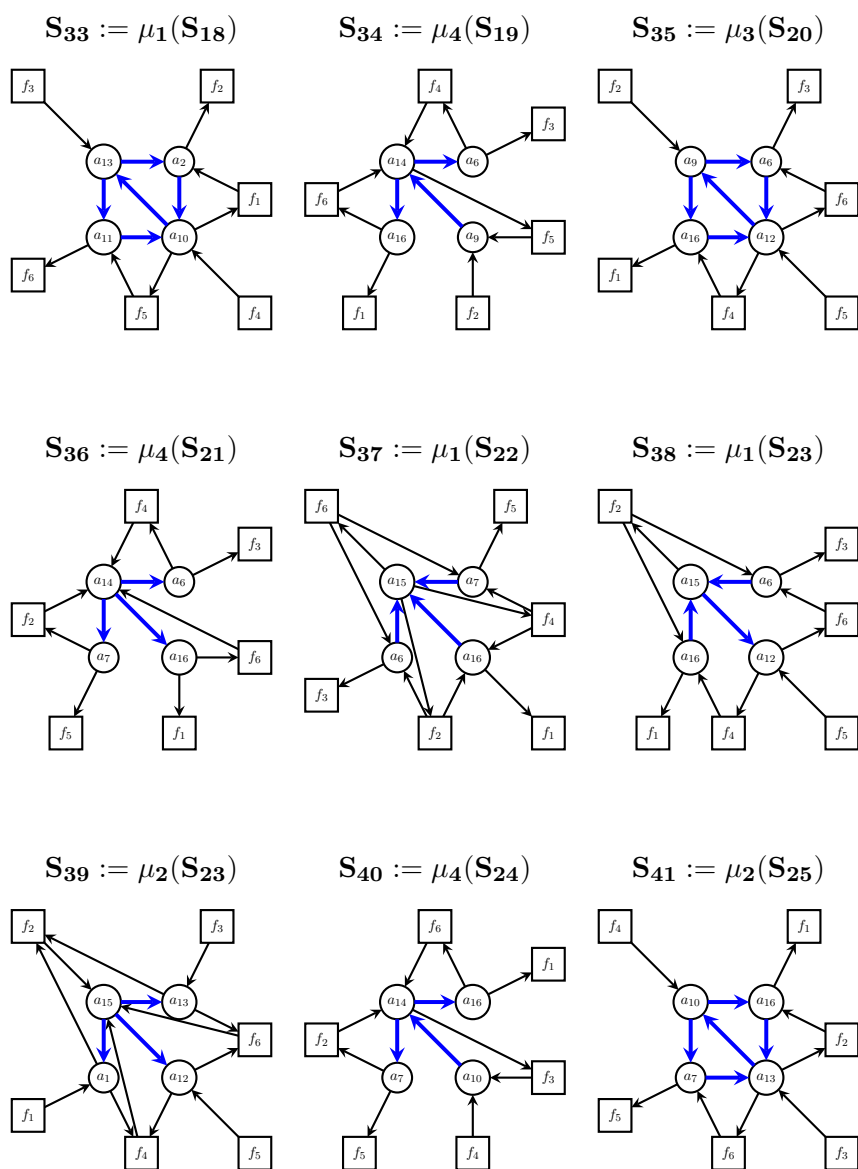


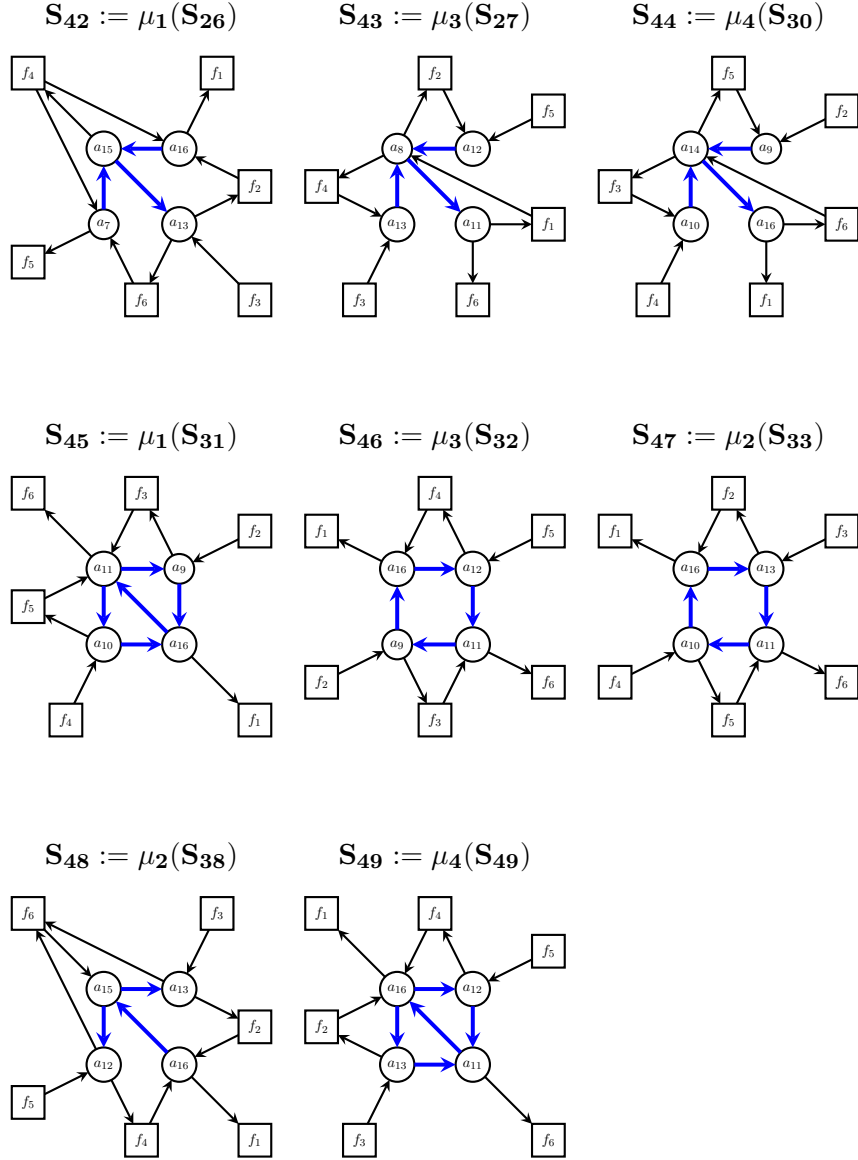
$$S_{31} := \mu_4(S_{16})$$



$$S_{32} := \mu_1(S_{17})$$







On a donc obtenu 50 graines différentes par mutations successives sur la graine S_0 . Il est intéressant de vérifier que toute autre combinaison de mutations appliquée à n'importe quelle de ces graines aura pour résultat une graine isomorphe à l'une ou l'autre de celles-ci. On a donc que l'anneau de coordonnées de la grassmannienne $\text{Gr}(3,6)$ est doté d'une structure d'algèbre amassée d'ordre finie. En effet, en observant le graphe de la graine S_1 , entre autres, on note que le sous-graphe composé des sommets mutables et des flèches les reliant (en bleu) est un graphe de Dynkin de type D_4 . Ceci vérifie donc le théorème 2.13.

Soit $\mathcal{A} := \{a_1, a_2, \dots, a_{16}, f_1, f_2, \dots, f_6\}$, l'algèbre amassée générée par la grai-

ne S_0 .

De plus, les mutations appliquée lors de la construction de ces graines permet de déterminer les relations d'échange suivantes :

$$\begin{aligned} a_1 \mathbf{a}_5 &= a_4 f_1 + a_2 a_3 & a_2 \mathbf{a}_6 &= a_1 f_3 + a_4 f_2 & a_3 \mathbf{a}_7 &= a_1 f_5 + a_4 f_4 \\ a_4 \mathbf{a}_8 &= a_1 f_3 f_5 + a_2 a_3 f_6 & a_2 \mathbf{a}_9 &= a_5 f_2 + f_1 f_3 & a_3 \mathbf{a}_{10} &= a_5 f_4 + f_1 f_5 \\ a_4 \mathbf{a}_{11} &= a_5 f_6 + f_3 f_5 & a_4 \mathbf{a}_{12} &= a_3 f_6 + a_6 f_5 & a_4 \mathbf{a}_{13} &= a_2 f_6 + a_7 f_3 \\ a_3 \mathbf{a}_{14} &= a_4 a_9 f_4 + a_6 f_1 f_6 & a_3 \mathbf{a}_{15} &= a_1 a_{12} + a_6 f_4 & a_3 \mathbf{a}_{16} &= a_9 f_4 + a_{12} f_1 \end{aligned}$$

De ces relations, on peut vérifier que chacune des variables de cette algèbre amassée peut être représentée par un polynôme de Laurent dont les termes sont composés des variables de l'amas initial x_0 :

$$\begin{aligned} a_5 &= \frac{a_4 f_1 + a_2 a_3}{a_1} & a_6 &= \frac{a_1 f_3 + a_4 f_2}{a_2} \\ a_7 &= \frac{a_1 f_5 + a_4 f_4}{a_3} & a_8 &= \frac{a_1 f_3 f_5 + a_2 a_3 f_6}{a_4} \\ a_9 &= \frac{a_1 f_1 f_3 + a_4 f_1 f_4 + a_2 a_3 f_2}{a_1 a_2} & a_{10} &= \frac{a_1 f_1 f_5 + a_4 f_1 f_4 + a_2 a_3 f_4}{a_1 a_3} \\ a_{11} &= \frac{a_1 f_3 f_5 + a_4 f_1 f_6 + a_2 a_3 f_6}{a_1 a_4} & a_{12} &= \frac{a_1 f_3 f_5 + a_4 f_2 f_5 + a_2 a_3 f_6}{a_2 a_4} \\ a_{13} &= \frac{a_1 f_3 f_5 + a_4 f_3 f_4 + a_2 a_3 f_6}{a_3 a_4} \\ a_{14} &= \frac{a_1 a_4 f_1 f_3 f_4 + a_1 a_4 f_1 f_2 f_5 + a_2 a_3 a_4 f_2 f_4 + a_1^2 f_1 f_3 f_5 + a_4^2 f_1 f_2 f_4}{a_1 a_2 a_3} \\ a_{15} &= \frac{a_1 a_4 f_3 f_4 + a_1 a_4 f_2 f_5 + a_1 a_2 a_3 f_6 + a_1^2 f_3 f_5 + a_4^2 f_2 f_4}{a_2 a_3 a_4} \\ a_{16} &= \frac{a_1 a_4 f_1 f_3 f_4 + a_1 a_4 f_1 f_2 f_5 + a_1 a_2 a_3 f_1 f_6 + a_2 a_3 a_4 f_2 f_4 + a_1^2 f_1 f_3 f_5 + a_4^2 f_1 f_2 f_4}{a_1 a_2 a_3 a_4} \end{aligned}$$

On peut donc constater que le phénomène de Laurent est vérifié dans le cas de la grassmanienne $\text{Gr}(3,6)$.

Ensuite, si on remplace respectivement les variables de l'amas initial,

$$x_0 = \{a_1, a_2, a_3, a_4, f_1, f_2, f_3, f_4, f_5, f_6\},$$

par un ensemble de variables de Plücker,

$$P_0 = \{X_{256}, X_{356}, X_{125}, X_{235}, X_{156}, X_{456}, X_{345}, X_{126}, X_{123}, X_{234}\},$$

dans les relations d'échanges obtenues précédemment, alors on obtient les nouvelles relations suivantes :

$$\begin{aligned} X_{256} \mathbf{a}_5 &= X_{235} X_{156} + X_{356} X_{125} & X_{356} \mathbf{a}_6 &= X_{256} X_{345} + X_{235} X_{456} \\ X_{125} \mathbf{a}_7 &= X_{256} X_{123} + X_{235} X_{126} & X_{235} \mathbf{a}_8 &= X_{256} X_{345} X_{123} + X_{356} X_{125} X_{234} \\ X_{356} \mathbf{a}_9 &= a_5 X_{456} + X_{156} X_{345} & X_{125} \mathbf{a}_{10} &= a_5 X_{126} + X_{156} X_{123} \end{aligned}$$

$$\begin{aligned}
X_{235}\mathbf{a}_{11} &= a_5 X_{234} + X_{345} X_{123} & X_{235}\mathbf{a}_{12} &= X_{125} X_{234} + a_6 X_{123} \\
X_{235}\mathbf{a}_{13} &= X_{356} X_{234} + a_7 X_{345} & X_{125}\mathbf{a}_{14} &= X_{235} a_9 X_{126} + a_6 X_{156} X_{234} \\
X_{125}\mathbf{a}_{15} &= X_{256} a_{12} + a_6 X_{126} & X_{125}\mathbf{a}_{16} &= a_9 X_{126} + a_{12} X_{156}
\end{aligned}$$

En utilisant la courte relation de Plücker, on peut maintenant facilement associer une variable de Plücker à chacune des variables amassées de \mathcal{A} à l'exception de a_8 et a_{14} :

<i>V. amassée</i>	<i>V. Plücker</i>	<i>V. amassée</i>	<i>V. Plücker</i>
f_1	X_{156}	f_2	X_{456}
f_3	X_{345}	f_4	X_{126}
f_5	X_{123}	f_6	X_{234}
a_1	X_{256}	a_2	X_{356}
a_3	X_{125}	a_4	X_{235}
a_5	X_{135}	a_6	X_{245}
a_7	X_{236}	a_9	X_{145}
a_{10}	X_{136}	a_{11}	X_{134}
a_{12}	X_{124}	a_{13}	X_{346}
a_{15}	X_{246}	a_{16}	X_{146}

Enfin, à l'aide de manipulations algébriques élémentaires sur les variables amassées ou par manipulations en utilisant la relation de Plücker, on peut déterminer que les variables a_8 et a_{14} sont en fait une combinaison linéaire des autres variables :

$$\begin{aligned}
a_8 &= a_2 a_{12} - f_2 f_5 = X_{124} X_{356} - X_{123} X_{456} \\
a_{14} &= a_7 a_9 - f_2 f_5 = X_{145} X_{236} - X_{123} X_{456}
\end{aligned}$$

Nous avons donc, comme l'indiquait le théorème 3.4, plus de variables amassées que de variables de Plücker. Dans la section suivante, cet exemple nous permettra également d'observer la validité d'un autre théorème qui nous vient de la théorie de représentation.

4 Les algèbres inclinées amassées

Buan, Marsh, Reineke, Reiten et Todorov ont développé la théorie des catégories amassées [BMR⁺06] et des algèbres inclinées amassées qui reproduit la combinatoire des variables amassées par les objets inclinés amassés [BMR08].

Après avoir vu ce qu'est une algèbre inclinée et une algèbre inclinée amassée, le but de cette section sera tout d'abord de définir et apprendre à calculer la

dimension globale d'une algèbre. Afin de caractériser la dimension globale d'une algèbre inclinée et d'une algèbre inclinée amassée.

Nous terminerons en décrivant les carquois des algèbres inclinées amassées reliés à $\text{Gr}(3,6)$ comme illustration.

4.1 Notations et remarques

Notation 4.1. Pour la suite, k est un corps algébriquement clos.

Notation 4.2. On notera par $\text{Hom}_C(O, O')$ la classe des morphismes entre les deux objets O et O' de la catégorie C , et $\text{End}_C O = \text{Hom}_C(O, O)$.

Notation 4.3. On notera par $\text{rep } Q$ la catégorie ayant pour objets les représentations finies du carquois Q . C'est-à-dire toutes les réalisations du carquois Q dans la catégorie des k espaces vectoriels. Il s'agit de remplacer tous les sommets du carquois par des k espaces vectoriels et toutes les flèches par des applications k linéaires [ASS06].

Remarque 4.4. À noter que la catégorie $\text{rep } Q$ est équivalente à la catégorie des modules de type fini sur l'algèbre de chemins kQ .

Pour conclure ces remarques ; dans la suite, on aura besoin du bifoncteur Ext_C^n (voir [ASS06], Appendix, page 426, ou [Ass97], pour une définition).

4.2 Définitions

On commence par quelques définitions indispensables, où n est le nombre des sommets du carquois Q .

Définition 4.5. Un objet T de la catégorie $\text{rep } Q$ est *inclinant* si

1. $\text{Ext}_{\text{rep } Q}^1(T, T) = 0$
2. $|T| = n$, c'est-à-dire $T = T_1 \oplus T_2 \oplus \cdots \oplus T_n$ où les T_i sont deux à deux non isomorphes et indécomposables.

On notera par $\mathcal{C}(Q)$ la catégorie amassée du carquois Q qui est le quotient de la catégorie bornée dérivée \mathcal{D} des modules de type fini sur une algèbre héréditaire, par le foncteur $F = \tau^{-1}[1]$ où τ est la translation d'Auslander-Reiten et $[1]$ est le foncteur translation. Cette catégorie a été définie par Buan, Marsh, Reineke, Reiten et Todorov. L'idée principale qu'ils ont amené était d'élargir la catégorie $\text{rep } Q$. Le point important est que dans une telle catégorie on va pouvoir définir un objet inclinant de la même manière que dans la catégorie $\text{rep } Q$, ce qui mène au résultat principal dans [BMR⁺06].

Voir [BMR⁺06] pour plus de détails.

Définition 4.6. Un objet T de la catégorie amassée $\mathcal{C}(Q)$ est *inclinant-amassé* si

- 1) $\text{Ext}_{\mathcal{C}(Q)}^1(T, T) = 0$
- 2) $|T| = n$, c'est-à-dire $T = T_1 \oplus T_2 \oplus \cdots \oplus T_n$ où les T_i sont deux à deux non isomorphes et indécomposables.

Définition 4.7. Si T est un objet inclinant dans $\text{rep } Q$ alors $A = \text{End}(T, T) = \text{Hom}_{\text{rep } Q}(T, T)$ est une algèbre *inclinée*.

Définition 4.8. Si T est inclinant amassé alors $B = \text{End}(T) = \text{Hom}_{\mathcal{C}(Q)}(T, T)$ est une algèbre *inclinée amassée*.

Maintenant dans le but de définir la dimension globale d'une algèbre, on définit la dimension projective d'un module :

Définition 4.9. Soit M un module sur une algèbre A . Le plus petit entier n tel qu'il existe une résolution projective

$$0 \rightarrow P_n \rightarrow P_{n-1} \rightarrow \cdots \rightarrow P_1 \rightarrow P_0 \rightarrow M \rightarrow 0$$

est dite *la dimension projective* de M : $\text{dp } M = n$.

Si une telle résolution finie n'existe pas, on a $\text{dp } M = \infty$ et $\text{dp } 0 = -\infty$.

Exemple 4.10. Pour tout module projectif P on a $\text{dp } P = 0$.

D'une manière duale, on peut définir la dimension injective $\text{di } M$ d'un module M (voir [ASS06], Appendix, page 426, (b) ou [Ass97]).

Théorème 4.11 (voir [ASS06], Appendix A.4).

$$\text{dp } M = \sup\{n \mid \text{Ext}_A^n(M, -) \neq 0\}$$

Puis, voilà la définition de la dimension globale d'une algèbre :

Définition 4.12. Le plus grand entier n (s'il existe) tel qu'il existe deux A -modules M et N tel que $\text{Ext}_A^n(M, N) \neq 0$ est appelé *la dimension globale* de l'algèbre A : $\text{dim. gl. } A = n$.

Si un tel entier n'existe pas on dit que $\text{dim. gl. } A = \infty$ et $\text{dim. gl. } 0 = -\infty$

Théorème 4.13 ([ASS06], Appendix A.4.). *Il existe une équivalence entre :*

1. $\text{dim. gl. } A \leq n$
2. $\text{dp } M \leq n$, pour tout M
3. $\text{di } M \leq n$, pour tout M .

Corollaire 4.14.

$$\text{dim. gl. } A = \sup\{\text{dp } M\} = \sup\{\text{di } M\}$$

Exemple 4.15. Une algèbre A est semi simple si et seulement si $\text{dim. gl. } A = 0$.

Théorème 4.16 (Auslander). (*Théorème 2.5, [Ass97] page 281*)

$$\text{dim. gl. } A = \sup\{\text{dp } A/I \mid I \text{ idéal de } A\}$$

Corollaire 4.17.

$$\dim. \text{gl. } A = \sup\{\text{dp } M \mid M \text{ de type fini}\}$$

Théorème 4.18. *Si l'algèbre A est artinienne alors*

$$\dim. \text{gl. } A = \sup\{\text{dp } S \mid S \text{ module simple}\}$$

Cette simplification est vraiment utile étant donné qu'il existe un nombre fini de A -modules simples non isomorphes sur une algèbre artinienne. Les calculs deviennent donc faisables.

Maintenant que nous avons tout le nécessaire, énonçons le théorème suivant caractérisant la dimension globale des algèbres inclinées :

Théorème 4.19 (Happel, Ringel ([HR82])). *On considère l'algèbre héréditaire KQ . Soient T_{KQ} un objet incliné et $A = \text{End}_{KQ} T$ une algèbre inclinée.*

Alors $\dim. \text{gl. } A \leq 2$ et pour tout A -module indécomposable Z , on a $\text{dp } Z \leq 1$ et $\text{di } Z \leq 1$.

Passons au cas inclinée amassée :

Théorème 4.20 (Keller, Reiten ([KR07])). *Une algèbre inclinée amassée B est 1-Gorenstein, c'est-à-dire que $\text{dp } I \leq 1$ pour tout A -module injectif I et $\text{di } P \leq 1$ pour tout A -module projectif P .*

Corollaire 4.21 (KR). *Tout B -module M est soit de dimension projective infinie, soit tel que $\text{dp } M \leq 1$.*

Preuve : Supposons que $\text{dp } M = n < \infty$ et montrons que $n \leq 1$.

On notera $\Omega M = \Omega^1 M$ le noyau (nommé le premier syzygy de M) de l'épimorphisme

$$0 \rightarrow \Omega M \rightarrow P \rightarrow M \rightarrow 0$$

où P est projectif.

Comme on a supposé que $\text{dp } M = n$ alors $\Omega^n M$ est projectif

$$0 \rightarrow \Omega^n M = P_n \rightarrow P_{n-1} \rightarrow \dots \rightarrow P_2 \rightarrow P_1 \rightarrow P \rightarrow M \rightarrow 0$$

Comme $\Omega^n M$ est un B -module projectif et l'algèbre B est Gorenstein, alors par le théorème 4.20 $\text{di } (\Omega^n M) \leq 1$; on la suite exacte courte

$$0 \rightarrow \Omega^n M \rightarrow P_{n-1} \rightarrow \Omega^{n-1} \rightarrow 0$$

$$\text{di } (\Omega^{n-1}) \leq \sup\{\text{di } \Omega^n - 1, \text{di } P_{n-1}\} = \sup\{0, 1\} = 1 \leq 1.$$

Par récurrence sur $p \leq 0$, on trouve que Ω^{n-p} est de $\text{di } \Omega^{n-p} \leq 1$.

D'où pour $p = n$

$$\text{di } (\omega^0 M) = \text{di } M \leq 1 \leq \infty.$$

De façon duale, on suppose que $\text{di } M \leq \infty$ et on obtient que $\text{dp } M \leq 1$. □

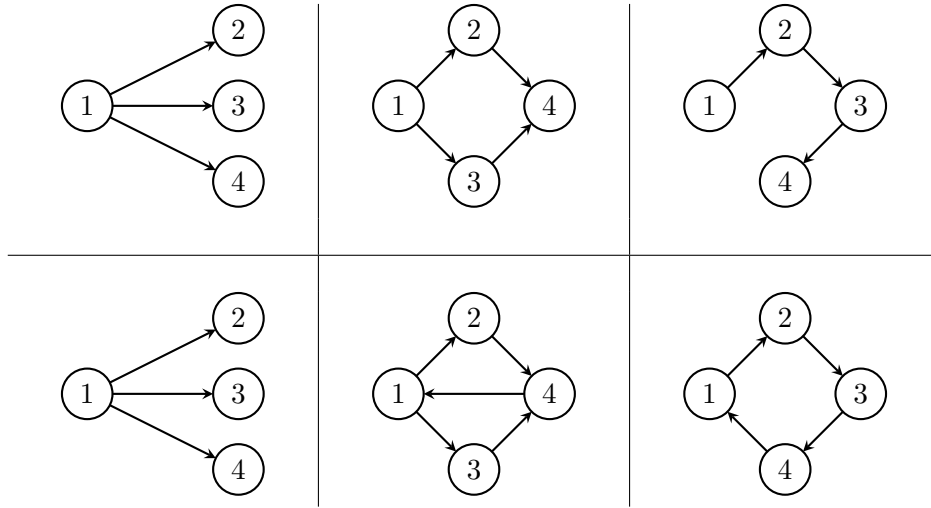
De même se démontre le corollaire suivant :

Corollaire 4.22. *Tout B -module est soit de dimension injectif infinie, soit tel que $\dim M \leq 1$.*

Par le théorème 4.13, ceci caractérise clairement la dimension globale d'une algèbre inclinée amassée.

Illustrons les caractérisations précédentes par des exemples :

Exemple 4.23. Motivés par le fait que la grassmannienne $\text{Gr}(3,6)$ est finie de type D_4 , nous avons pris le carquois de type D_4 et nous avons calculé tous les objets inclinants et tous les objets inclinants amassés de type D_4 (c'est-à-dire les objets dans la catégorie $\text{rep } Q$ pour $Q = D_4$). Ensuite, nous avons calculés leurs anneaux d'endomorphismes. De cette manière, on a obtenu toutes les algèbres inclinées et inclinées amassées de type D_4 . Dans le tableau ci-dessous, nous présentons leurs carquois (Les algèbres de la première ligne sont des algèbres inclinées alors que celles de la deuxième ligne sont des algèbres inclinées amassées.) :



Elles ont pour dimensions globales respectives : $\dim. \text{gl. } A = 1, 2$ et 2 alors que $\dim. \text{gl. } B = 1, \infty$ et ∞ , ce qui vérifie clairement les théorème de caractérisation déjà énoncés plus haut.

On observe le miracle déjà prédit par Buan, Marsh, et Reiten[BMR08] : ces carquois de la deuxième ligne sont exactement les carquois bleus des graines du type $\text{Gr}(3,6)$ qu'on a déjà calculé ; voir les diagrammes de la sous-section 3.1.2.

Remarquons que les algèbres inclinées amassées possèdent plus de flèches que celles inclinées. Ceci est dû au fait qu'on forme notre anneau d'endomorphismes dans une catégorie plus large.

Références

[Ass97] Ibrahim ASSEM : *Algèbres et modules*. Masson, Paris, 1997.

- [ASS06] Ibrahim ASSEM, Daniel SIMSON et Andrzej SKOWROŃSKI : *Elements of the representation theory of associative algebras. Vol. 1*, volume 65 de *London Mathematical Society Student Texts*. Cambridge University Press, Cambridge, 2006. Techniques of representation theory.
- [BMR⁺06] Aslak Bakke BUAN, Robert MARSH, Markus REINEKE, Idun REITEN et Gordana TODOROV : Tilting theory and cluster combinatorics. *Adv. Math.*, 204(2):572–618, 2006.
- [BMR08] Aslak Bakke BUAN, Robert J. MARSH et Idun REITEN : Cluster mutation via quiver representations. *Comment. Math. Helv.*, 83(1): 143–177, 2008.
- [FZ02] Sergey FOMIN et Andrei ZELEVINSKY : Cluster algebras. I. Foundations. *J. Amer. Math. Soc.*, 15(2):497–529, 2002.
- [HR82] Dieter HAPPEL et Claus Michael RINGEL : Tilted algebras. *Trans. Amer. Math. Soc.*, 274(2):399–443, 1982.
- [KR07] Bernhard KELLER et Idun REITEN : Cluster-tilted algebras are Gorenstein and stably Calabi-Yau. *Adv. Math.*, 211(1):123–151, 2007.

SOUHEILA HASSOUN

DÉPARTEMENT DE MATHÉMATIQUES, UNIVERSITÉ DE SHERBROOKE

Courriel: Souheila.Hassoun@USherbrooke.ca

DENIS LANGFORD

DÉPARTEMENT DE MATHÉMATIQUES, UNIVERSITÉ DE SHERBROOKE

Courriel: Denis.Langford@USherbrooke.ca

FRÉDÉRIC LANGLOIS

DÉPARTEMENT DE MATHÉMATIQUES, UNIVERSITÉ DE SHERBROOKE

Courriel: Frédéric.Langlois@USherbrooke.ca

Dégénération des représentations des carquois

Vincent Roy

Reçu le 2017-10-05 et accepté le 2017-11-11

RÉSUMÉ Cet article traite de la dégénération des représentations des carquois. Nous allons en premier lieu définir quelques aspects topologiques et d'actions de groupes. Ensuite, nous introduisons le concept de représentation de carquois et appliquons ce qui a été défini plus tôt. Finalement, nous définissons la dégénération sur un carquois de type \mathbb{A}_2 .

1 Introduction

Les représentations de carquois sont une notion faisant partie de la théorie des représentations qui fait l'étude de structures algébriques. Plus précisément, la notion de représentation de carquois est utile lorsqu'on s'intéresse à des problèmes d'algèbre linéaire, par exemple pour visualiser un ensemble algébrique par l'entremise d'un ensemble de K -espaces vectoriels.

Dans cet article, il est question de carquois de type \mathbb{A}_n et plus spécifiquement des carquois de type \mathbb{A}_2 . Pour bien comprendre, des notions de topologie sont utiles comme la fermeture d'un ensemble [Har77]. Aussi, la définition de certaines applications [Sim83] est nécessaire en vue d'utiliser certains principes d'actions de groupes.

2 Topologie de Zariski

On pose K comme étant un corps algébrique clos tout au long de cet article. Soit K muni d'un n -espace affine noté \mathbb{A}^n . On définit celui-ci comme étant un ensemble de n -uplets d'éléments de K . Un élément de \mathbb{A}^n est appelé un *point* et est de la forme $p = (p_1, \dots, p_n)$ tel que $p_i \in K$, $1 \leq i \leq n$, où les p_i sont appelés les *coordonnées* de p .

On définit maintenant $K[x_1, \dots, x_n]$ comme étant un anneau de polynômes à n variables sur K . Ensuite, on associe chaque polynôme $f(x_1, \dots, x_n)$ à une application de \mathbb{A}^n vers K en définissant $f(p) = f(p_1, \dots, p_n)$, pour tout point

Je tiens à grandement remercier le SAG pour son financement ainsi que M. Shiping Liu professeur à l'Université de Sherbrooke pour la supervision de mon stage de recherche et pour la précieuse aide qu'il m'a offerte pour la rédaction de cet article.

$p = (p_1, \dots, p_n) \in \mathbb{A}^n$. Ainsi, on peut définir l'ensemble des zéros de f , qu'on nomme $Z(f) = \{p \in \mathbb{A}^n \mid f(p) = 0\}$. De façon générale, soit T un sous-ensemble de A , on définit l'ensemble des zéros de T :

$$Z(T) = \{p \in \mathbb{A}^n \mid f(p) = 0, \forall f \in T\}.$$

Définition 2.1. Un sous-ensemble Y de \mathbb{A}^n est un *ensemble algébrique* s'il existe T , un sous-ensemble de $K[x_1, \dots, x_n]$, tel que $Y = Z(T)$.

Proposition 2.2. *L'union de deux ensembles algébriques est un ensemble algébrique. L'intersection de n'importe quelle famille d'ensembles algébriques est un ensemble algébrique. De plus, l'ensemble vide et l'espace au complet sont des ensembles algébriques.*

Démonstration. Soient deux ensembles algébriques $Y_1 = Z(T_1)$ et $Y_2 = Z(T_2)$. Alors, $Y_1 \cup Y_2 = Z(T_1 T_2)$, où $T_1 T_2$ représente l'ensemble de tous les produits d'un élément de T_1 par un élément de T_2 . Donc, si on prend $p \in Y_1 \cup Y_2$, on a que soit $p \in Y_1$ ou $p \in Y_2$, alors p est un zéro de tous les polynômes dans $T_1 T_2$. Ainsi, on peut affirmer que $Y_1 \cup Y_2 \subseteq Z(T_1 T_2)$.

Inversement, si on prend $p \in Z(T_1 T_2)$ et on suppose que $p \notin Y_1$, on a qu'il existe un $f \in T_1$ tel que $f(p) \neq 0$. Par le fait que $p \in Z(T_1 T_2)$, on doit avoir, pour tout $g \in T_2$, l'égalité $(fg)(p) = 0$, ce qui implique que $g(p) = 0$, donc $p \in Y_2$. Ainsi, on peut affirmer que $Z(T_1 T_2) \subseteq Y_1 \cup Y_2$.

Ainsi, l'union de deux ensembles algébriques est un ensemble algébrique, car $Y_1 \cup Y_2 = Z(T_1 T_2)$.

Si $Y_i = Z(T_i)_{i \in I}$ est une famille d'ensembles algébriques, on a l'égalité $\bigcap_{i \in I} Y_i = Z(\bigcup_{i \in I} T_i)$, alors $\bigcap_{i \in I} Y_i$ est un ensemble algébrique.

Finalement, l'ensemble vide est un ensemble algébrique, car $\emptyset = Z(1)$ et l'espace au complet est aussi un ensemble algébrique. car $\mathbb{A}^n = Z(0)$. \square

Définition 2.3. La *topologie de Zariski* sur \mathbb{A}^n est définie par les sous-ensembles ouverts complémentaires aux ensembles algébriques.

Ceci est bien une topologie puisque, par la proposition 2.2, on a la définition d'un espace topologique par les fermés. Donc, en prenant les complémentaires des fermés, on obtient la définition d'un espace topologique par les ouverts.

Définition 2.4. Soit X un espace topologique et soit A un sous-ensemble de X . On définit la *fermeture* de A , notée \overline{A} , comme le plus petit ensemble fermé contenant A .

Remarque 2.5. En d'autres mots, \overline{A} est l'intersection de tous les ensembles fermés contenant A . Ainsi, il est évident que si A est fermé, alors on a que $A = \overline{A}$.

Proposition 2.6. *Soit X un espace topologique. Si A et B sont des sous-ensembles de X , alors l'opération de fermeture respecte les quatre propriétés suivantes :*

1. $\overline{\emptyset} = \emptyset$
2. $A \subseteq \overline{A}$
3. $\overline{\overline{A}} = \overline{A}$
4. $\overline{A \cup B} = \overline{A} \cup \overline{B}$.

Lemme 2.7. *Soit U un sous-ensemble propre de \mathbb{A}^n . Alors, U est fermé si et seulement si U est fini.*

Démonstration. Posons d'abord $U = \{P_1\}$, avec $P_1 = (p_1, \dots, p_n)$. On considère les polynômes $f_i \in K[x_1, \dots, x_n]$ de la forme $f_i = (x_i - p_i)$, $1 \leq i \leq n$. Ainsi, on obtient que $\{P_1\} = Z(f_1, \dots, f_n)$. Donc, par définition, $\{P_1\}$ est fermé. De façon plus générale, prenons maintenant $U = \{P_1, \dots, P_t\}$, tel que U est un sous-ensemble fini quelconque de \mathbb{A}^n . On peut écrire $U = \bigcup_{i=1}^t \{P_i\}$ comme une union finie d'ensembles fermés. Or, on sait par la proposition 2.2 que l'union d'un nombre fini d'ensembles fermés est aussi un ensemble fermé. Donc, U est un ensemble fermé.

Supposons que U est fermé. On a que $U = Z(T)$, où $T \subset K[x_1, \dots, x_n]$ est non-vide. Soit un polynôme $f \in T$ tel que $f(x_1, \dots, x_n) = (x_1 - p_1) \cdots (x_n - p_n)$ où $p_1, \dots, p_n \in K$. Donc, $Z(f) = \{p \in \mathbb{A}^n \mid f(p) = 0\}$. Alors, on a que $U \subseteq Z(f)$, car $U = Z(T) = \bigcap_{g \in T} Z(g) \subseteq Z(f)$. Ainsi, U est fini. \square

Proposition 2.8. *Soit U un sous-ensemble propre de \mathbb{A}^1 . Alors U est fermé si et seulement si U est fini.*

Démonstration. La suffisance découle du lemme 2.7.

Supposons que U est fermé. Alors, $U = Z(T)$ où $T \subset K[x]$ non-vide. Prenons $f \in T$. Comme K est algébriquement clos, $f(x) = (x - p_1) \cdots (x - p_n)$ où $p_1, \dots, p_n \in K$. Donc, $Z(f) = \{p_1, \dots, p_n\}$ et on a que $U \subseteq Z(f)$, car $U = Z(T) = \bigcap_{g \in T} Z(g) \subseteq Z(f)$. Donc, U est fini. \square

Remarque 2.9. Par conséquent, si on prive \mathbb{A}^n d'un sous-ensemble fini Y , alors l'ensemble $\mathbb{A}^n \setminus Y$ est un ensemble ouvert.

Définition 2.10. Une application $\phi : \mathbb{A}^n \rightarrow \mathbb{A}^m$ est dite *régulière* s'il existe ϕ_1, \dots, ϕ_m in $K[x_1, \dots, x_n]$ tel que $\phi(p) = (\phi_1(p), \dots, \phi_m(p))$, pour tout $p \in \mathbb{A}^n$.

Définition 2.11. Soient une application $\varphi : \mathbb{A}^n \rightarrow \mathbb{A}^m$ et un sous-ensemble V de \mathbb{A}^m . On définit $\varphi^{-1}(V) = \{p \in \mathbb{A}^n \mid \varphi(p) \in V\}$, un sous-ensemble de \mathbb{A}^n .

Définition 2.12. Une application $\varphi : \mathbb{A}^n \rightarrow \mathbb{A}^m$ est dite *continue* si, pour tout sous-ensemble fermé V de \mathbb{A}^m , l'ensemble $\varphi^{-1}(V)$ est fermé dans \mathbb{A}^n .

Lemme 2.13. *Si une application $\phi : \mathbb{A}^n \rightarrow \mathbb{A}^m$ est régulière, alors elle est continue.*

Démonstration. Par hypothèse, l'application ϕ est régulière, alors, par la définition 2.10, il existe $\phi_1, \dots, \phi_m \in K[x_1, \dots, x_n]$ tels que $\phi(p) = (\phi_1(p), \dots, \phi_m(p))$, pour tout $p \in \mathbb{A}^n$. Prenons U un ensemble fermé de \mathbb{A}^m . Alors, il existe un ensemble $T \subseteq K[y_1, \dots, y_m]$ tel que $U = Z(T)$. Pour tout $f \in T$, on pose

$$g_f = f(\phi_1(x_1, \dots, x_n), \dots, \phi_m(x_1, \dots, x_n)) \in K[x_1, \dots, x_n].$$

On obtient donc l'ensemble $S = \{g_f \mid f \in T\}$, on prétend que $\phi^{-1}(U) = Z(S)$, où $Z(S) \subseteq \mathbb{A}^n$. En effet, pour tout $p \in \mathbb{A}^n$, on a que $p \in \phi^{-1}(U)$ si et seulement si $\phi(p) = (\phi_1(p), \dots, \phi_m(p)) \in U$, si et seulement si $f(\phi(p)) = 0$, pour tout $f \in T$. Or, on a par définition que $f(\phi(p)) = g_f(p)$. Donc, pour tout $f \in T$, $f(\phi(p)) = 0$ si et seulement si $g_f(p) = 0$, pour tout $g_f \in S$, c'est à dire $p \in Z(S)$. Par conséquent, $\phi^{-1}(U) = Z(S)$, qui est fermé par définition. Donc, par la définition d'une application continue, on peut affirmer que ϕ est continue. \square

Définition 2.14. Un espace topologique X est dit *irréductible* si X ne peut pas s'écrire de la forme $X = X_1 \cup X_2$ où X_1 et X_2 sont deux sous-ensembles propres et fermés de X .

Lemme 2.15. Soit X un espace topologique irréductible. Si U est un sous-ensemble ouvert non vide de X alors $\overline{U} = X$.

Démonstration. Par la proposition 2.2, on sait que l'espace complet est fermé. Ainsi, X est fermé et donc $\overline{X} = X$. Par définition, on a $X = U \cup (X \setminus U)$. En prenant la fermeture de chaque côté de l'égalité et par la proposition 2.6, on obtient $X = \overline{U} \cup (X \setminus U)$. Or, on a que \overline{U} est un ensemble fermé et $X \setminus U$ est un ensemble fermé propre de X . Donc, on doit avoir que $X = \overline{U}$, car X est irréductible. \square

Proposition 2.16. \mathbb{A}^1 est un espace topologique irréductible.

Démonstration. Supposons au contraire que \mathbb{A}^1 n'est pas irréductible. Alors, on peut écrire $\mathbb{A}^1 = Y_1 \cup Y_2$, où Y_1 et Y_2 sont deux sous-ensembles fermés et propres de \mathbb{A}^1 . Or, on sait par le lemme 2.7 que Y_1 et Y_2 sont des ensembles finis. Ainsi $Y_1 \cup Y_2$ est aussi un ensemble fini. Par conséquent, \mathbb{A}^1 est fini, mais \mathbb{A}^1 est infini, car K est un corps algébriquement clos. Ceci mène à une contradiction. \square

En utilisant la définition de la fermeture d'un ensemble et la définition d'application continue, il est possible de déterminer si un élément est dans la fermeture d'un ensemble.

Lemme 2.17. Soit $U \subseteq \mathbb{A}^n$ et $f : \mathbb{A}^1 \rightarrow \mathbb{A}^n$ une application régulière. Si $f(x) \in U$ pour tout $x \in K^* = \mathbb{A}^1 \setminus \{0\}$, alors $f(0) \in \overline{U}$.

Démonstration. Du fait que \overline{U} est un ensemble fermé et que f est continue par le lemme 2.13, on sait par définition que $f^{-1}(\overline{U})$ est fermé dans \mathbb{A}^1 . Par hypothèse, on a que $f(x) \in U$ pour tout $x \in K^*$. Donc, $K^* \subset f^{-1}(\overline{U})$. Comme K^* est un ensemble ouvert et \mathbb{A}^1 est irréductible, d'après le lemme 2.15, on a que $\overline{K^*} = \mathbb{A}^1$. De plus, comme $f^{-1}(\overline{U})$ est fermé, $\overline{K^*} \subseteq f^{-1}(\overline{U})$. En particulier, $0 \in f^{-1}(\overline{U})$, c'est-à-dire $f(0) \in \overline{U}$. \square

3 Les actions de groupes

Définition 3.1. Soient X un espace topologique et G un groupe. On dit que G agit sur X s'il existe une opération :

$$\begin{aligned} G \times X &\rightarrow X \\ (g, x) &\mapsto g \cdot x \end{aligned}$$

telle que

1. $e \cdot x = x$ pour tout $x \in X$, où e est l'identité de G ;
2. $g_1 \cdot (g_2 \cdot x) = (g_1 \cdot g_2) \cdot x$ pour tout $g_1, g_2 \in G$ et pour tout $x \in X$;
3. Pour tout $g \in G$, l'application

$$\begin{aligned} g : X &\rightarrow X \\ x &\mapsto g \cdot x \end{aligned}$$

est un homéomorphisme ou, en d'autres mots, un isomorphisme entre deux espaces topologiques, ce qui implique que l'application g est bijective et que g et g^{-1} sont continues.

En posant G , un groupe agissant sur X , on définit, pour tout $x \in X$, l'ensemble $G \cdot x = \{g \cdot x \mid g \in G\}$ qu'on nomme la G -orbite de x .

Lemme 3.2. Soit G un groupe agissant sur X . Pour tout $Y \subseteq X$ et pour tout $g \in G$, on a que $g \cdot \overline{Y} = \overline{g \cdot Y}$.

Démonstration. Prenons une application qui est un homéomorphisme défini comme :

$$\begin{aligned} g : X &\rightarrow X \\ x &\mapsto g \cdot x. \end{aligned}$$

Ainsi, on obtient que $g \cdot \overline{Y}$ est fermé, car \overline{Y} est fermé et g^{-1} est continue. De plus, $g \cdot Y \subseteq g \cdot \overline{Y}$ puisque $Y \subseteq \overline{Y}$. De ce fait, on peut affirmer que $\overline{g \cdot Y} \subseteq g \cdot \overline{Y}$, pour tout $g \in G$ et pour tout $Y \subseteq X$.

Dans l'autre sens, prenons $g^{-1} \in G$ et $g \cdot Y \subseteq X$. Ainsi, comme vu précédemment, on a que $\overline{Y} = \overline{g^{-1} \cdot (g \cdot Y)} \subseteq g^{-1} \cdot \overline{g \cdot Y}$. En appliquant g de chaque côté, on obtient $g \cdot \overline{Y} \subseteq \overline{g \cdot Y}$. Donc $g \cdot \overline{Y} = \overline{g \cdot Y}$. \square

Théorème 3.3. Soit G un groupe agissant sur X et soit $x \in X$. Si $y \in \overline{G \cdot x}$, alors $G \cdot y \subseteq \overline{G \cdot x}$.

Démonstration. Pour tout $g \in G$, on a que $g \cdot y \in g \cdot \overline{G \cdot x}$. Or, par le lemme 3.2, on a que $g \cdot \overline{G \cdot x} = \overline{g \cdot (G \cdot x)}$. Cependant, $g \cdot (G \cdot x) = G \cdot x$, car $g \in G$. Ainsi, $G \cdot y \subseteq \overline{G \cdot x}$. \square

4 Représentations des carquois

Définition 4.1. Un *carquois* Q est un graphe orienté qui se compose d'un ensemble Q_0 de sommets et d'un ensemble Q_1 de flèches entre les sommets. Si $\alpha : a \rightarrow b$ est une flèche, alors a s'appelle la *source* de α , notée $s(\alpha)$, et b , le *but* de α , noté $b(\alpha)$.

Définition 4.2. Un carquois $Q = (Q_0, Q_1)$ est dit *fini* si Q_0 et Q_1 sont des ensembles finis.

Définition 4.3. Soit $Q = (Q_0, Q_1)$ un carquois fini. Une *représentation* K -linéaire M de Q se compose d'une famille d'espaces vectoriels de dimension finie sur K , notée $(M(a))_{a \in Q_0}$, et d'une famille d'applications K -linéaires, notée $(M(\alpha))_{\alpha \in Q_1}$, telles que, pour une flèche $\alpha : a \mapsto b$, on a $M(\alpha) : M(a) \rightarrow M(b)$.

Définition 4.4. Soit $Q = (Q_0, Q_1)$ un carquois fini avec $Q_0 = \{1, 2, \dots, n\}$. Un *vecteur-dimension* est un vecteur $\underline{d} = (d_i)_{i \in Q_0}$ d'entiers non négatifs. En outre, \underline{d} est dit *sincère* si $d_i > 0$ pour tout $i \in Q_0$.

Remarque 4.5. En particulier, si M est une représentation de Q , alors le vecteur-dimension de M est un vecteur $\underline{d} = (d_i)_{i \in Q_0}$, où $d_i = \dim(M(i))$ pour tout $i \in Q_0$.

Définition 4.6. Soient Q un carquois fini et \underline{d} un vecteur-dimension. On note $\text{Rep}_K(Q, \underline{d})$ l'ensemble des représentations K -linéaires de Q de vecteur-dimension \underline{d} .

Dans ce qui suit, si $f : U \rightarrow V$ et $g : V \rightarrow W$ sont des applications linéaires de K -espaces vectoriels, alors l'application composée de f et g est notée $f \circ g$.

Définition 4.7. Soient $Q = (Q_0, Q_1)$ un carquois fini et M, N deux représentations K -linéaires de Q . Un *homomorphisme* $f : M \rightarrow N$ est composé d'une famille d'applications K -linéaires $f(a) : M(a) \rightarrow N(a)$ avec $a \in Q_0$ telles que, pour tout $\alpha \in Q_1$, le diagramme suivant commute :

$$\begin{array}{ccc} M(a) & \xrightarrow{f(a)} & N(a) \\ M(\alpha) \downarrow & & \downarrow N(\alpha) \\ M(b) & \xrightarrow{f(b)} & N(b), \end{array}$$

ce qui veut dire que $f(a) \circ N(\alpha) = M(\alpha) \circ f(b)$. On dit que $f : M \rightarrow N$ est un *isomorphisme* si, pour tout $a \in Q_0$, $f(a)$ est un isomorphisme d'espaces vectoriels sur K .

Par exemple, si on considère le carquois \mathbb{A}_n suivant pour $n \geq 2$:

$$1 \xrightarrow{\alpha_1} 2 \xrightarrow{\alpha_2} \dots \xrightarrow{\alpha_{n-2}} n-1 \xrightarrow{\alpha_{n-1}} n,$$

alors une représentation M du carquois \mathbb{A}_n sur le corps K peut être illustrée par le diagramme suivant :

$$M(1) \xrightarrow{M(\alpha_1)} M(2) \longrightarrow \dots \longrightarrow M(n-1) \xrightarrow{M(\alpha_{n-1})} M(n).$$

Un homomorphisme $f : M \rightarrow N$ peut être illustré par le diagramme commutatif suivant :

$$\begin{array}{ccccccc} M(1) & \xrightarrow{M(\alpha_1)} & M(2) & \longrightarrow & \dots & \longrightarrow & M(n-1) \xrightarrow{M(\alpha_{n-1})} M(n) \\ f(1) \downarrow & & f(2) \downarrow & & & & f(n-1) \downarrow & & f(n) \downarrow \\ N(1) & \xrightarrow{N(\alpha_1)} & N(2) & \longrightarrow & \dots & \longrightarrow & N(n-1) \xrightarrow{N(\alpha_{n-1})} N(n), \end{array}$$

où $f(i)$ est une application K -linéaire.

Définition 4.8. Soient $Q = (Q_0, Q_1)$ un carquois fini avec $Q_0 = \{1, 2, \dots, n\}$ et \underline{d} un vecteur-dimension sincère. Une K -linéaire *représentation matricielle* de Q de vecteur-dimension $\underline{d} = (d_i)_{i \in Q_0}$ est composée d'une famille $\{M_\alpha \mid \alpha \in Q_1\}$ de matrices sur K telle que, pour toute flèche $\alpha : i \rightarrow j$, la matrice M_α est de la taille $d_i \times d_j$.

Soient Q un carquois fini et \underline{d} un vecteur-dimension sincère. L'ensemble des K -linéaires représentations matricielles de Q de vecteur-dimension \underline{d} est noté $\text{Mat}(Q, \underline{d})$.

Soient $Q = (Q_0, Q_1)$ un carquois fini avec $Q_0 = \{1, 2, \dots, n\}$ et un vecteur-dimension $\underline{d} = (d_i)_{i \in Q_0}$. On pose :

$$\mathcal{GL}(K, \underline{d}) = \prod_{i \in Q_0} \mathcal{GL}(K, d_i),$$

où $\mathcal{GL}(K, d_i)$ est le groupe des matrices inversibles sur K d'ordre d_i . Alors, pour toutes représentations matricielles $(M_\alpha)_{\alpha \in Q_1} \in \text{Mat}(Q, \underline{d})$ et pour tout $g = (g_i)_{i \in Q_0} \in \mathcal{GL}(K, \underline{d})$, on pose :

$$g \cdot (M_\alpha)_{\alpha \in Q_1} = (g_{s(\alpha)}^{-1} \cdot M_\alpha \cdot g_{b(\alpha)})_{\alpha \in Q_1}.$$

Proposition 4.9. *L'action définie ci-dessus est une action de $\mathcal{GL}(K, \underline{d})$ sur $\text{Mat}(Q, \underline{d})$.*

Démonstration. On doit montrer que cette action respecte les trois axiomes de la définition 3.1. La preuve des axiomes 1 et 2 est évidente.

Pour l'axiome 3, on doit avoir un homéomorphisme, donc il faut montrer que l'action est continue. Pour ce faire, montrons que pour tout $g = (g_i)_{i \in Q_0}$ l'application suivante est une application régulière :

$$\begin{array}{ccc} g : \text{Mat}(Q, \underline{d}) & \rightarrow & \text{Mat}(Q, \underline{d}) \\ M_\alpha & \mapsto & (g_{s(\alpha)}^{-1} \cdot M_\alpha \cdot g_{b(\alpha)})_{\alpha \in Q_1}. \end{array}$$

Pour $\alpha \in Q_1$, posons $d_{s(\alpha)} = m$ et $d_{b(\alpha)} = n$. Ainsi, on a que $g_{s(\alpha)}^{-1} \in \mathcal{GL}(K, m)$, $g_{b(\alpha)} \in \mathcal{GL}(K, n)$ et $M_\alpha \in M_{m \times n}(K)$. Soient $g_{s(\alpha)}^{-1} = [a_{ij}]$ pour $1 \leq i, j \leq m$, $g_{b(\alpha)} = [b_{ij}]$ pour $1 \leq i, j \leq n$ et $M_\alpha = [x_{ij}]$ pour $1 \leq i \leq m$ et $1 \leq j \leq n$. Donc, si on pose la matrice $g_{s(\alpha)}^{-1} \cdot M_\alpha \cdot g_{b(\alpha)} = [z_{ij}]$, on a $z_{ij} = \sum_{l=1}^n (\sum_{k=1}^m a_{ik} \cdot b_{lj}) \cdot x_{kl}$. Défini de cette façon, on a que g est une application régulière. Alors, g est continue par le lemme 2.13. \square

Ainsi, si le carquois Q est du type \mathbb{A}_n avec $n \geq 2$, on a $|Q_0| = n$, $|Q_1| = n-1$ et un vecteur-dimension $\underline{d} = (d_1, \dots, d_n)$ de la représentation $M = (M_1, \dots, M_{n-1})$ où $M_i \in M_{d_i \times d_{i+1}}(K)$ pour $i = 1, \dots, n-1$, alors le groupe qui agit sur $\text{Mat}(\mathbb{A}_n, \underline{d})$ est

$$\mathcal{GL}(K, \underline{d}) = \mathcal{GL}(K, d_1) \times \mathcal{GL}(K, d_2) \times \dots \times \mathcal{GL}(K, d_n).$$

Posons les deux variables suivantes, $x = (g_1, g_2, \dots, g_n) \cdot (M_1, M_2, \dots, M_{n-1})$ et $y = (g_1^{-1} \cdot M_1 \cdot g_2, g_2^{-1} \cdot M_2 \cdot g_3, \dots, g_{n-1}^{-1} \cdot M_{n-1} \cdot g_n)$, où $g_i \in \mathcal{GL}(K, d_i)$, pour tout $i = 1, 2, \dots, n$. Alors, l'action de groupe est définie par $x = y$.

Définition 4.10. Soient Q un carquois avec $Q_0 = \{1, 2, \dots, n\}$, $\underline{d} = (d_1, \dots, d_n)$ un vecteur-dimension et une représentation matricielle $M \in \text{Mat}(\mathbb{A}_n, \underline{d})$, on définit l'orbite de M , notée $\mathcal{O}(M)$ comme suit :

$$\{g \cdot M \mid g \in \mathcal{GL}(K, \underline{d})\}.$$

Pour introduire le prochain théorème, on construit une correspondance entre $\text{Mat}(Q, \underline{d})$ et $\text{Rep}_K(Q, \underline{d})$.

Soient un carquois fini $Q = (Q_0, Q_1)$, un vecteur-dimension sincère \underline{d} et une représentation M de $\text{Rep}_K(Q, \underline{d})$. Pour tout $i \in Q_0$, on fixe une base \mathcal{B}_i de $M(i)$. Pour chaque flèche $\alpha : i \rightarrow j$, on prend M_α comme étant la matrice de $M(\alpha) : M(i) \rightarrow M(j)$ dans les bases \mathcal{B}_i et \mathcal{B}_j . On obtient alors une famille de matrices $\{M_\alpha \mid \alpha \in Q_1\}$ qui est une représentation matricielle de vecteur-dimension \underline{d} de Q .

Dans l'autre sens, étant donnée une représentation matricielle $\{M_\alpha \mid \alpha \in Q_1\}$ de $\text{Mat}(Q, \underline{d})$, pour tout $i \in Q_0$, on pose $M(i) = K^{d_i}$ et, pour tout $\alpha : i \rightarrow j$, on pose :

$$\begin{aligned} M(\alpha) : K^{d_i} &\rightarrow K^{d_j} \\ u &\mapsto u \cdot M_\alpha. \end{aligned}$$

On obtient une représentation M de Q .

Théorème 4.11. Soient $Q = (Q_0, Q_1)$ un carquois et un vecteur-dimension sincère \underline{d} . Alors, les classes d'isomorphismes de $\text{Rep}(Q, \underline{d})$ sont en bijection avec les $\mathcal{GL}(K, n)$ -orbites de $\text{Mat}(Q, \underline{d})$.

Démonstration. Soit une application qui associe chaque classe d'isomorphisme de $\text{Rep}(Q, \underline{d})$ à une orbite de $\text{Mat}(Q, \underline{d})$. Vérifions si l'application est correctement définie. Prenons une représentation M et fixons, pour chaque espace

vectorel $M(a)$, une base $\mathcal{B}(a)$. Pour chaque application linéaire $M(\alpha)$, on a une matrice $M_\alpha = [M(\alpha)]_{\mathcal{B}(b)}^{\mathcal{B}(a)}$ et, ainsi, on obtient une famille $\{M_\alpha \mid \alpha \in Q_1\}$. Alors, on envoie la représentation M sur l'orbite $\mathcal{O}(\{M_\alpha \mid \alpha \in Q_1\})$. Fixons une autre base $\mathcal{B}'(a)$ pour chaque $M(a)$. On obtient donc une nouvelle famille $\{M'_\alpha \mid \alpha \in Q_1\}$ telle que $M'_\alpha = [M(\alpha)]_{\mathcal{B}'(b)}^{\mathcal{B}'(a)}$. Or, il existe $g \in \mathcal{GL}(K, \underline{d})$ tel que $g \cdot \{M_\alpha \mid \alpha \in Q_1\} = \{M'_\alpha \mid \alpha \in Q_1\}$. Alors, on a que $\{M'_\alpha \mid \alpha \in Q_1\}$ est dans $\mathcal{O}(\{M_\alpha \mid \alpha \in Q_1\})$ et les deux orbites coïncident. Donc, l'application est correctement définie.

Maintenant, montrons la bijection. Soient $M, N \in \text{Rep}(Q, \underline{d})$ telles que les ensembles $\{M_\alpha \mid \alpha \in Q_1\}$ et $\{N_\alpha \mid \alpha \in Q_1\}$ sont leurs représentations matricielles respectives. Supposons que $\mathcal{O}(\{M_\alpha \mid \alpha \in Q_1\}) = \mathcal{O}(\{N_\alpha \mid \alpha \in Q_1\})$. Ainsi, il existe $g \in \mathcal{GL}(K, \underline{d})$ tel que $g \cdot \{M_\alpha \mid \alpha \in Q_1\} = \{N_\alpha \mid \alpha \in Q_1\}$ avec $g = (g_1, g_2, \dots, g_n)$ où $g_i \in \mathcal{GL}(K, d_i)$, pour tout $i = 1, 2, \dots, n$. On peut illustrer cela par le diagramme commutatif suivant :

$$\begin{array}{ccccccc} M(1) & \xrightarrow{M_{\alpha_1}} & M(2) & \longrightarrow & \dots & \longrightarrow & M(n-1) \xrightarrow{M_{\alpha_{n-1}}} M(n) \\ g_1 \downarrow & & g_2 \downarrow & & & & g_{n-1} \downarrow & & g_n \downarrow \\ N(1) & \xrightarrow{N_{\alpha_1}} & N(2) & \longrightarrow & \dots & \longrightarrow & N(n-1) \xrightarrow{N_{\alpha_{n-1}}} N(n). \end{array}$$

On a n isomorphismes $g_i : M(i) \rightarrow N(i)$ entre les représentations M et N . Donc, M et N sont isomorphes et l'application est injective. De plus, on sait que c'est surjectif, car, pour chaque orbite de $\text{Mat}(Q, \underline{d})$, on a au moins une représentation de $\text{Rep}(Q, \underline{d})$. \square

Remarque 4.12. $\text{Mat}(Q, \underline{d})$ est un sous-espace topologique de $\mathbb{A}^{d_1 \cdots d_n}$.

Définition 4.13. Soient $Q = (Q_0, Q_1)$ un carquois fini, \underline{d} un vecteur-dimension de Q et deux représentations $M_1, M_2 \in \text{Mat}(Q, \underline{d})$. On dit que M_2 est une *dégénération* de M_1 (ou que M_1 se *dégénère* en M_2), notée $M_1 \leq_{deg} M_2$, si $\mathcal{O}(M_2) \subseteq \overline{\mathcal{O}(M_1)}$. De plus, on dit que M_2 est une *dégénération stricte* de M_1 (ou M_1 se *dégénère strictement* en M_2), notée $M_1 <_{deg} M_2$, si $\mathcal{O}(M_2) \subset \overline{\mathcal{O}(M_1)}$.

5 Carquois \mathbb{A}_2

Dans cette section, nous travaillons sur un carquois du type \mathbb{A}_2 , en d'autres mots, un carquois composé de deux sommets et une flèche. Ainsi, une représentation de \mathbb{A}_2 est composée de deux espaces vectoriels et une application linéaire. Un vecteur-dimension de ce carquois est de la forme $\underline{d} = (d_1, d_2)$. Une représentation matricielle de \mathbb{A}_2 est un élément de $M_{m \times n}(K)$. Ainsi, $\text{Mat}(\mathbb{A}_2, \underline{d}) = M_{m \times n}(K)$. Nous notons cet espace topologique \mathbb{A}^{mn} . L'anneau des coordonnées de \mathbb{A}^{mn} est $K[x_{ij} \mid 1 \leq i \leq m; 1 \leq j \leq n]$. Posons $G = \mathcal{GL}(K, m) \times \mathcal{GL}(K, n)$.

Soient $A, B \in M_{m \times n}(K)$. On voit que A et B sont dans la même G -orbite si et seulement si il existe deux matrices $g_1 \in \mathcal{GL}(K, m)$ et $g_2 \in \mathcal{GL}(K, n)$ telles que $Ag_2 = g_1B$.

Proposition 5.1. *Si $A \in M_{m \times n}(K)$, alors la G -orbite de A est définie par $\mathcal{O}(A) = \{B \in M_{m \times n}(K) \mid rg(B) = rg(A)\}$.*

Démonstration. Supposons $B \in \mathcal{O}(A)$. Alors, on a que $B = g_1^{-1}Ag_2$. Donc, $rg(B) = rg(A)$. Supposons $rg(B) = rg(A) = r$. On peut réduire A à la matrice échelonnée suivante :

$$I_r = \begin{pmatrix} 1 & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & \ddots & & & & \vdots \\ \vdots & & 1 & 0 & & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & & & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 \end{pmatrix}.$$

Alors, il existe $g_1 \in \mathcal{GL}(K, m)$ et $g_2 \in \mathcal{GL}(K, n)$ tels que $I_r = g_1Ag_2$ ou que $I_r = g'_1Bg'_2$. Donc, $B \in \mathcal{O}(A)$. \square

Lemme 5.2. *Soit un nombre entier r tel que $0 \leq r \leq \min(m, n)$. L'ensemble $F_r = \{A \in M_{m \times n}(K) \mid rg(A) \leq r\}$ est fermé.*

Démonstration. Tout d'abord, si $r = \min(m, n)$, alors $F_r = M_{m \times n}(K)$ est fermé par la proposition 2.2.

Maintenant, posons $r < \min(m, n)$ et considérons la matrice d'inconnus suivante :

$$X = \begin{pmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & & \vdots \\ x_{m1} & \cdots & x_{mn} \end{pmatrix}.$$

Prenons X_1, \dots, X_p , les sous-matrices carrées d'ordre $r+1$ de X . En prenant le déterminant de X_k , on obtient un polynôme $m_k \in K[x_{ij} \mid 1 \leq i \leq m; 1 \leq j \leq n]$ où $1 \leq k \leq p$. Soit $A \in M_{m \times n}(K)$. Alors, $A \in F_r$ si et seulement si $rg(A) \leq r$, si et seulement si les mineurs d'ordre $r+1$ de A sont tous nuls, si et seulement si A est annulée par les polynômes m_1, \dots, m_p , c'est-à-dire que $F_r = Z(m_1, \dots, m_p)$, qui est fermé par définition. \square

Proposition 5.3. *Soit $A \in M_{m \times n}(K)$. On a que*

$$\overline{\mathcal{O}(A)} = \{B \in M_{m \times n}(K) \mid rg(B) \leq rg(A)\}.$$

Démonstration. Soit $rg(A) = r$. En particulier, $A \in F_r$. On a que $\mathcal{O}(A) \subseteq F_r$. Étant donné que $\overline{\mathcal{O}(A)}$ est le plus petit fermé contenant $\mathcal{O}(A)$, d'après le lemme 5.2, on a que $\overline{\mathcal{O}(A)} \subseteq F_r$.

Maintenant, supposons $B \in F_r$, telle que $rg(B) = s$, avec $s \leq r$. Montrons par récurrence que $B \in \overline{\mathcal{O}(A)}$. Supposons $s = r$, d'après la proposition 5.1, on a que $B \in \mathcal{O}(A)$ et, donc, $B \in \overline{\mathcal{O}(A)}$. Supposons que cela est vrai pour $s \leq r$. Par

des opérations élémentaires, il est possible d'obtenir, à partir de B , la matrice échelonnée $E_s \in M_{m \times n}(K)$ suivante :

$$E_s = \begin{pmatrix} I_s & 0 \\ 0 & 0 \end{pmatrix}.$$

Soit l'application continue suivante :

$$\begin{aligned} f : K &\rightarrow M_{m \times n}(K) \\ \lambda &\mapsto \begin{pmatrix} E_{s-1} & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & 0 \end{pmatrix}. \end{aligned}$$

Par hypothèse de récurrence, $f(\lambda) \in \overline{\mathcal{O}(A)}$ pour tout $\lambda \in K^*$. Par le lemme 2.17, on a que $f(0) \in \overline{\mathcal{O}(A)}$. Or, $f(0) = E_{s-1}$, alors $E_{s-1} \in \overline{\mathcal{O}(A)}$. Si $rg(B) = s - 1$, alors B s'échelonne à E_{s-1} et on obtient que $B \in \mathcal{O}(E_{s-1}) \subseteq \overline{\mathcal{O}(A)}$. Donc, $\overline{\mathcal{O}(A)} = \{B \in M_{m \times n}(K) \mid rg(B) \leq rg(A)\}$. \square

Théorème 5.4. *Soient $A, B \in \text{Mat}(\mathbb{A}_2, \underline{d})$. Alors, $A \leq_{deg} B$ si et seulement si $rg(B) \leq rg(A)$.*

Démonstration. Par définition, on a que $A \leq_{deg} B$ si et seulement si $B \in \overline{\mathcal{O}(A)}$. En vertu de la proposition 5.3, cette condition est équivalente à $rg(B) \leq rg(A)$. \square

Références

- [Har77] Robin HARTSHORNE : *Graduate Texts in Mathematics : Algebraic Geometry*. Springer-Verlag, New-York, 1977.
- [Sim83] George F. SIMMONS : *Introduction to Topology and Modern Analysis*. Krieger Publishing Company, Malabar, Florida, 1983.

VINCENT ROY
 DÉPARTEMENT DE MATHÉMATIQUES, UNIVERSITÉ DE SHERBROOKE
 Courriel: `vincent.Z.roy@USherbrooke.ca`

Analyse fine d’algorithmes de type ARC-TR au cas unidimensionnel

Samuel Goyette

Reçu le 2017-09-21 et accepté le 2018-01-09

RÉSUMÉ Dans ce document, nous présenterons les algorithmes d’optimisation unidimensionnelle de région de confiance et de régularisation avec adaptation cubique. Puis, nous étudierons le calcul de minimiseurs globaux pour chacun d’entre eux. Nous verrons également des résultats numériques pour les algorithmes de régularisation cubique.

1 Introduction

Le problème d’optimisation que l’on veut résoudre est le suivant

$$\underset{t}{\text{minimiser}} \phi(t)$$

où $\phi : \mathbb{R} \rightarrow \mathbb{R}$ est une fonction continue deux fois différentiable. S’il existe un point t^* qui résout ce problème, il doit satisfaire aux conditions nécessaires d’optimalité :

$$\phi'(t^*) = 0 \tag{1}$$

$$\phi''(t^*) \geq 0. \tag{2}$$

Ainsi, pour résoudre ce problème, on peut utiliser une méthode analytique : on dérive la fonction, on trouve les points qui annulent la dérivée (appelés des points stationnaires) et, parmi ces points, on regarde les points dont l’évaluation de la dérivée seconde est positive. Cependant, certaines fonctions sont compliquées et le calcul n’est pas trivial. Si on prend les polynômes de degré un ou deux, il existe des formules analytiques connues pour trouver les racines de tels polynômes. Il en existe aussi pour ceux de degrés trois et quatre, mais elles sont complexes et peuvent mener à des erreurs de précision numérique. Pour les polynômes d’ordre supérieur ou égal à cinq, il n’existe pas de formule analytique permettant de trouver les racines. C’est le cas de nombreuses fonctions.

J’aimerais remercier Jean-Pierre Dussault qui m’a supervisé dans toutes mes activités de recherche au cours de mon baccalauréat.

On utilise donc des algorithmes d'optimisation qui nous permettent de trouver les racines de fonctions.

Les algorithmes de *région de confiance* (appelés TR pour *Trust-Regions*) et les algorithmes de *régularisation avec adaptation cubique* (appelés ARC pour *Adaptive Regularization using Cubics algorithms*) ont pour idée générale d'approximer une fonction "complexe" par une fonction plus "simple", que l'on appelle modèle, sur un certain intervalle. Ces fonctions sont dites "simples", car on est facilement capable de calculer leurs points stationnaires (ainsi que leurs minima globaux), ce qui permet de minimiser la fonction originale rapidement.

Dans ce texte, nous allons étudier les algorithmes de région de confiance et les algorithmes de régularisation cubique adaptative dans les sections 2 et 3. Dans ces deux sections, nous observerons le fonctionnement traditionnel des algorithmes puis certaines modifications qui permettent de les simplifier. Dans la section 4, nous observerons les résultats numériques associés aux deux méthodes.

2 Algorithme région de confiance (TR)

2.1 Présentation de l'algorithme

Les algorithmes TR visent à approximer notre fonction initiale ϕ par une fonction quadratique q_t . On veut minimiser cette approximation dans un certain intervalle $[-\Delta, \Delta]$ où $\Delta > 0$ que l'on appelle région de confiance. Supposons que l'on a ϕ deux fois différentiable (remarquons que certaines variantes n'utilisent qu'une approximation de la dérivée seconde, ce qui permet d'alléger cette hypothèse). On considère l'approximation suivante de ϕ dans un voisinage de t :

$$q_t(d) = \phi(t) + \phi'(t)d + \frac{1}{2}Hd^2 \approx \phi(t+d) \text{ pour } d \in [-\Delta, \Delta] \quad (3)$$

où H est soit $\phi''(t)$ soit une approximation de $\phi''(t)$.

L'idée est donc de minimiser q_t dans l'intervalle de confiance, afin de s'approcher du minimum de ϕ . Il s'agit d'un procédé itératif, c'est-à-dire qu'après s'être approché du minimum de ϕ , on ajuste q_t pour avoir la meilleure approximation possible de ϕ . On répète cette procédure jusqu'à l'obtention d'un minimum local de ϕ (ou un point qui s'en rapproche suffisamment). Chaque itération consiste à trouver d^* qui résout le problème suivant :

$$\begin{array}{ll} \underset{d}{\text{minimiser}} & q_t(d) \\ \text{sujet à} & |d| \leq \Delta \end{array}$$

Si q_t est une bonne approximation de ϕ dans $[-\Delta, \Delta]$, on peut se permettre d'augmenter l'intervalle de confiance, donc Δ . Si q_t est une mauvaise approximation, alors on réduit Δ . Pour déterminer si on doit ajuster la taille de l'intervalle de confiance, on compare la réduction prévue par le modèle et la réduction actuelle de la fonction que l'on note respectivement par $\text{pred} = q_t(0) - q_t(d_R)$ et $\text{ared} = \phi(t) - \phi(t + d_R)$. À l'aide du ratio de ces deux quantités $r = \frac{\text{ared}}{\text{pred}}$ on a

un aperçu de la qualité de l'approximation. On pose r_1 et r_2 des seuils pour qualifier q_t . Si on a un r trop petit ($r < r_1$), notre fonction q_t n'est pas une bonne approximation de ϕ dans $[-\Delta, \Delta]$, alors on réduit Δ et on passe à la prochaine itération. Sinon on ajuste la valeur de t . Finalement, après avoir ajusté la valeur de t , si r est assez grand ($r > r_2$), on a que q_t est une bonne approximation de ϕ sur $[-\Delta, \Delta]$ et on peut alors se permettre d'augmenter la taille de notre région de confiance. Nous verrons que le choix des paramètres r_1 et r_2 influence la performance de l'algorithme plus tard.

Le pseudo-code est présenté dans l'algorithme 1 [Dus15].

Algorithme 1 : Région de confiance modèle quadratique

Données : $\{\phi : \text{une fonction une ou deux fois différentiable}\};$
 $\{q_t \text{ une approximation quadratique telle que } q_t(0) = \phi(t)\};$
 $\{\epsilon \text{ un critère d'arrêt}\};$
 $\{r_1, r_2 \text{ des critères pour déterminer si on augmente la taille de l'intervalle de confiance}\};$

```

1  $\Delta \leftarrow 2 \quad t \leftarrow 1;$ 
2 répéter
3    $d_R \leftarrow \text{minimiseur global de } q_t \text{ dans l'intervalle } |d| \leq \Delta;$ 
4    $\text{ared} \leftarrow \phi(t) - \phi(t + d_R);$ 
5    $\text{pred} \leftarrow q_t(0) - q_t(d_R);$ 
6    $r \leftarrow \frac{\text{ared}}{\text{pred}};$ 
7   si ( $r < r_1$ ) alors  $\Delta \leftarrow \Delta/2;$ 
8   sinon
9      $t \leftarrow t + d_R;$ 
10    si  $r > r_2$  alors  $\Delta \leftarrow 2\Delta;$ 
11 jusqu'à ( $|\phi'(t)| \leq \epsilon$ );
12 Résultat  $\leftarrow t$ 

```

L'algorithme de région de confiance a plusieurs propriétés intéressantes :

- il est relativement simple ;
- il possède des propriétés de convergence présentées dans le théorème suivant.

Théorème 2.1. *Lorsque $H = \phi''(t)$, l'algorithme de région de confiance converge vers un point t^* qui satisfait aux conditions nécessaires d'optimalité (équations 1 et 2).*

La preuve de convergence de l'algorithme n'est pas triviale, pour plus de détail voir [Dus15]. Notons toutefois que la convergence est obtenue "à l'infini" (de manière asymptotique). Étant donné qu'en réalité on dispose d'un nombre fini d'itérations, ce minimum ne pourra jamais être atteint. Ceci explique pourquoi notre critère d'arrêt est $|\phi'(t)| < \epsilon$ et non $\phi'(t) = 0$.

2.2 Choix de H

Pour déterminer le minimum global de $q_t(d)$, on utilise différentes variantes du modèle 3. Ces variantes consistent simplement en différentes manières d'estimer la dérivée seconde de ϕ , c'est-à-dire, le coefficient quadratique de $q_t(d)$. Afin d'alléger la notation, posons $f = \phi(t)$ et $g = \phi'(t)$. On a

$$q_t(d) = f + gd + \frac{1}{2}Hd^2.$$

On peut utiliser les modèles suivants :

1. Le modèle de Taylor consiste à utiliser $H = \phi''(t)$. Dans ce cas, le polynôme de Taylor de degré 2 de $\phi(t+d)$ est $q_t(d) = \phi(t) + \phi'(t)d + \frac{1}{2}\phi''(t)d^2$. Dans la littérature, ainsi que dans cet article, on le note modèle de Newton.
2. On peut également utiliser le modèle de la sécante : $H = \frac{g-\tilde{g}}{t-\tilde{t}} \approx \phi''(t)$ où \tilde{t} est le t de l'itération précédente et $\tilde{g} = g(\tilde{t})$.
3. Un autre modèle possible est celui de la sécante améliorée. On pose : $s = t - \tilde{t}$, $y = g - \tilde{g}$, $\Gamma = 3(g - \tilde{g})s + 6(f - \tilde{f})$ et $\bar{y} = y + \frac{\Gamma}{s}$ où $\tilde{f} = f(\tilde{t})$. On a donc $H = \frac{\bar{y}}{s} \approx \phi''(t)$. Il s'agit d'une variante de la méthode de la sécante. Cette méthode est un cas particulier d'un algorithme en grande dimension, voir [ZX01] et [XZ01] pour plus de détail.

Dans tous les cas, le sommet de la parabole est approché par la direction de Newton $d_N = \frac{-g}{H}$. En effet,

$$q'_t(d) = g + Hd = 0 \Leftrightarrow d = \frac{-g}{H}.$$

On note que la variable par rapport à laquelle on dérive est d .

2.3 Choix de la direction de descente

Dans les méthodes de type région de confiance, afin de déterminer d (algorithme 1-ligne 4), on procède traditionnellement comme dans l'algorithme 2 en fonction du modèle considéré. Le d^* qui minimise $q_t(d)$ dans l'intervalle de confiance $[-\Delta, \Delta]$ représente la direction de descente.

Algorithme 2 : Choix traditionnel de la direction de descente pour les méthodes de région de confiance

```

1 Calcul de la direction de Newton :  $d_N = -g/H$ ;
2 si  $q_t(\Delta) < q_t(-\Delta)$  alors
3    $d_R = \Delta$ ;
4 sinon
5    $d = -\Delta$ ;
6 si  $|d_N| < \Delta$  et  $q_t(d) > q_t(d_N)$  alors  $d = d_N$ ;
7  $d_R \leftarrow d$ 
```

Le procédé est simple, on évalue la fonction q_t aux bornes de notre intervalle ainsi qu'à son sommet et la plus petite valeur représente la direction de descente. Or, on peut déterminer d_R sans jamais évaluer la fonction q_t grâce à une analyse simple d'un polynôme de degré deux. Le théorème suivant établit la direction de descente en fonction du signe de H et de g .

Théorème 2.2. *Si on a $H > 0$ (donc une parabole orientée vers le haut), alors la direction de descente d^* est soit d_N ou la borne de l'intervalle de confiance de signe opposé à g . Sinon, si $H < 0$ (donc une parabole orientée vers le bas), alors la direction de descente est la borne de l'intervalle de confiance de signe opposé à g .*

On peut visualiser les différentes possibilités décrites dans le théorème précédent dans la table 1. On peut également s'intéresser aux cas où g ou H sont nuls. Il s'agit de cas particuliers, car obtenir zéro en calcul numérique est un phénomène plutôt rare.

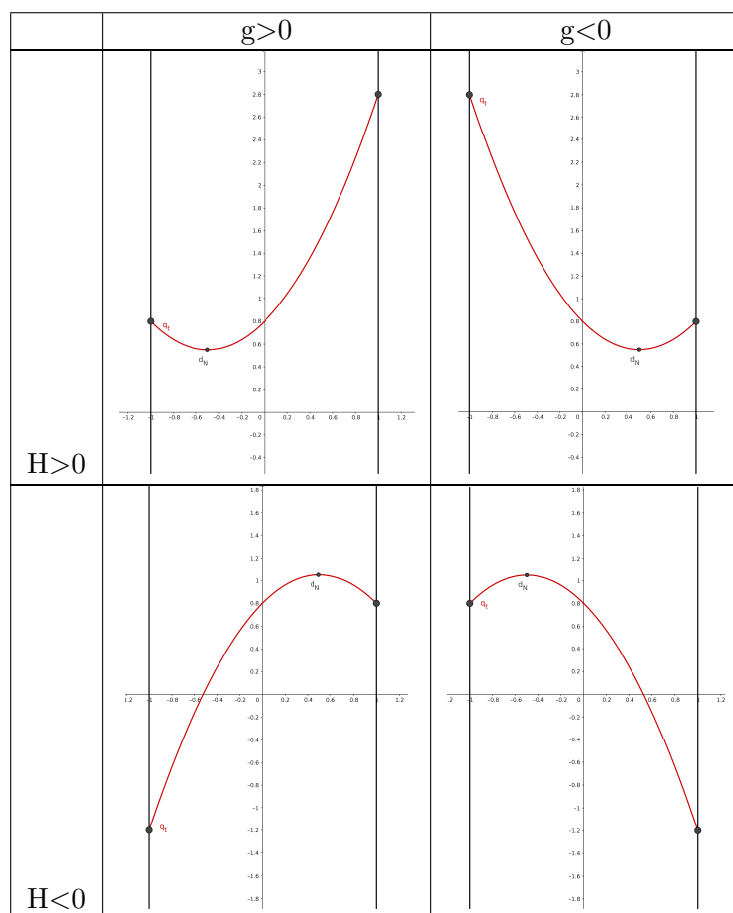
- **Cas 1 :** $g = 0$. Si $H > 0$, alors $d = 0$ est un minimum global de $q_t(d)$ donc l'algorithme arrête. Sinon si $H < 0$, on se déplace vers l'extrémité de l'intervalle qui se rapproche le plus du minimum local et on arrête, donc $d = \pm\Delta$.
- **Cas 2 :** $H = 0$. Dans ce cas $q_t(d)$ est une droite (qui a un minimum à l'intérieur de l'intervalle de confiance). Donc, il s'agit de la même procédure que lorsque $H < 0$.

On peut donc représenter ce théorème dans l'algorithme 3. On rappelle que cette procédure vise à trouver d tel que nécessaire à la ligne 4 de l'algorithme 1.

Algorithme 3 : Choix de la direction de descente pour TR	
<hr/>	
1	Calcul de la direction de Newton : $d_N = -g/H$;
2	si $H > 0$ alors
3	si $g > 0$ alors $d_R = \max(-\Delta, d_N)$;
4	sinon $d = \min(d_N, \Delta)$;
5	sinon
6	si $g > 0$ alors $d = -\Delta$;
7	sinon $d = \Delta$;
8	$d_R \leftarrow d$ (minimiseur global de q_t dans $[-\Delta, \Delta]$)

Changer l'algorithme 2 par l'algorithme 3 dans l'algorithme de région de confiance n'affecte en rien les performances de ce dernier. Il permet simplement des appels de moins à la fonction q_t que la version traditionnelle.

Table 1 – Cas possibles de région de confiance



Remarque 2.3. L'idée derrière la preuve mathématique de cet algorithme est très simple. On a que $q'_t(0) = g$, le signe de g détermine donc la variation de q_t à l'origine. On sait que q_t est représenté par une parabole donc le signe de H détermine si elle est ouverte vers le bas ($H < 0$) ou si elle est ouverte vers le haut ($H > 0$).

- **Cas 1 :** Si $H > 0$ et $g > 0$, alors q_t est croissante en 0. Donc l'abscisse de son sommet est négative, car q_t est ouverte vers le haut. Ainsi, on compare simplement $-\Delta$ et d_N . Si $-\Delta \leq d_N$, alors on prend d_N (qui correspond au sommet) sinon, d_N n'est pas dans l'intervalle de confiance donc on prend $-\Delta$. Si $g < 0$, il s'agit du même cas, mais avec une symétrie par rapport à la droite $x = 0$.
- **Cas 2 :** De la même manière, si $H < 0$, q_t est ouverte vers le bas. Donc le sommet n'est jamais un minimum local (c'est un maximum local). Ainsi le minimum dans l'intervalle correspond simplement au Δ le plus loin de d_N .

3 Algorithme de régularisation cubique adaptative (ARC)

3.1 Présentation de l'algorithme

L'algorithme ARC suit la même idée que l'algorithme de TR, mais en ajoutant un terme de régularisation cubique. Donc, on ne cherche plus à minimiser $q_t(d)$ (équation 3), mais bien :

$$c_t(d) = q_t(d) + \frac{1}{3\alpha} |d|^3 \text{ où } \alpha > 0. \quad (4)$$

Dans cet algorithme, la région de confiance disparaît (il n'y a plus de Δ , il n'y a plus de contrainte). Le terme α a le même rôle que Δ . Plus α est grand, plus on se rapproche de $q_t(d)$ et inversement, si α est petit, on a une fonction dont les valeurs augmentent rapidement lorsque d tend vers l'infini (négatif ou positif). Le pseudo-code est présenté dans l'algorithme 4.

Algorithme 4 : Algorithme ARC

Données : $\{\phi : \text{une fonction une ou deux fois dérivable}\};$
 $\{c_t \text{ une approximation de } \phi\};$
 $\{\epsilon \text{ un critère d'arrêt}\};$
 $\{r_1, r_2 \text{ des critères pour déterminer si on augmente la taille de l'intervalle de confiance}\};$

```

1   $t \leftarrow 0;$ 
2   $\tilde{t} \leftarrow 1;$ 
3  répéter
4       $d \leftarrow \text{minimiseur global de } c_t(d);$ 
5       $\text{ared} \leftarrow \phi(t) - \phi(t + d);$ 
6       $\text{pred} \leftarrow c_t(0) - c_t(d);$ 
7       $r \leftarrow \frac{\text{ared}}{\text{pred}};$ 
8      si  $(r < r_1)$  alors  $\alpha \leftarrow \alpha/2;$ 
9      sinon
10          $\tilde{t} \leftarrow t;$ 
11          $t \leftarrow t + d;$ 
12     si  $r > r_2$  alors  $\alpha \leftarrow 2\alpha;$ 
13 jusqu'à  $(|\phi'(t)| \leq \epsilon);$ 
14 Résultat  $\leftarrow t$ 

```

La forme de l'algorithme est très similaire à celle de l'algorithme de région de confiance. La différence entre les deux est que l'approximation de ϕ est maintenant c_t et non q_t . Étant donné qu'on a une autre approximation, la stratégie pour déterminer le minimiseur global de c_t n'est plus la même que celle utilisée pour la région de confiance. C'est ce que nous allons explorer dans la section qui suit.

3.2 Calcul du minimiseur global simple et précis

Le problème qui nous intéresse est de trouver la valeur de d qui minimise $c_t(d)$ (ligne 4 de l'algorithme 4). La difficulté supplémentaire qui vient avec la recherche d'un minimiseur de $c_t(d)$ réside dans le terme cubique avec une valeur absolue absent de la méthode TR. Il y a potentiellement deux minima locaux à cette fonction et il n'y a pas de formule directe pour nous trouver le minimum global. Cependant, il est tout de même possible d'énoncer un théorème similaire au théorème 2.1 pour l'algorithme ARC.

Théorème 3.1. *Soit $c_t(d) = f + gd + \frac{1}{2}Hd^2 + \frac{1}{3\alpha}|d|^3$. Posons $\delta_1 = H^2 - 4\frac{g}{\alpha}$ et $\delta_2 = H^2 + 4\frac{g}{\alpha}$. Si $H > 0$ alors le minimum global de c_t est :*

$$d^* = \frac{-2g}{H + \sqrt{\delta_1}}.$$

Sinon, si $H \leq 0$ alors :

$$d^* = \frac{-H + \sqrt{\delta_i}}{\frac{(-2 \operatorname{signe}(g))}{\alpha}}.$$

Où $i = 1$ si $g < 0$ et $i = 2$ si $g > 0$.

Démonstration. On a

$$c_t(d) = f + gd + \frac{1}{2}Hd^2 + \frac{1}{3\alpha}|d|^3.$$

On dérive c_t pour trouver les points stationnaires.

$$c'_t(d) = g + Hd + \frac{1}{\alpha}d|d|.$$

Maintenant, on cherche les racines de $c'_t(d)$. Étant donné qu'il y a une valeur absolue ainsi que plusieurs cas possibles en vertu des signes de g et H , on représente les racines possibles (ainsi que le nombre total de racines) dans la table 2.

Table 2 – Racines de $c'_t(d)$

	$d > 0$	$d < 0$	nombre de racines
$H > 0$ et $g > 0$	-	$\frac{-H + \sqrt{H^2 + 4\frac{g}{\alpha}}}{-\frac{2}{\alpha}}$	1
$H < 0$ et $g > 0$	$\frac{-H \pm \sqrt{H^2 - 4\frac{g}{\alpha}}}{\frac{2}{\alpha}}$	$\frac{-H + \sqrt{H^2 + 4\frac{g}{\alpha}}}{-\frac{2}{\alpha}}$	3
$H > 0$ et $g < 0$	$\frac{-H + \sqrt{H^2 - 4\frac{g}{\alpha}}}{\frac{2}{\alpha}}$	-	1
$H < 0$ et $g < 0$	$\frac{-H + \sqrt{H^2 - 4\frac{g}{\alpha}}}{\frac{2}{\alpha}}$	$\frac{-H \pm \sqrt{H^2 + 4\frac{g}{\alpha}}}{-\frac{2}{\alpha}}$	3

On peut aussi s'intéresser aux cas particuliers où $H = 0$ et $g = 0$. Les racines associées à ces cas sont dans la table 3.

Table 3 – Racines de $c'(d)$ (cas particuliers)

	$d > 0$	$d < 0$	nombre de racines
$H > 0$ et $g = 0$	0	0	1
$H < 0$ et $g = 0$	$\frac{-H}{\frac{1}{\alpha}}, 0$	$\frac{H}{\frac{1}{\alpha}}, 0$	2
$H = 0$ et $g > 0$	-	$\frac{\sqrt{4(\frac{g}{\alpha})}}{-\frac{2}{\alpha}}$	1
$H = 0$ et $g < 0$	$\frac{\sqrt{-4(\frac{g}{\alpha})}}{\frac{2}{\alpha}}$	-	1
$H = 0$ et $g = 0$	0	0	1

Maintenant que l'on a les racines de c'_t , on cherche à savoir lesquelles de celles-ci correspondent à des minima de c_t . Par construction de la méthode, lorsqu'on a une seule racine de $c'_t(d)$, il s'agit du minimum de c_t , car $\frac{1}{3\alpha}|d|^3 > 0$.

Lorsque $c'_t(d)$ a trois racines, les deux racines aux extrémités sont nécessairement des minima locaux de c_t et la racine au centre doit être un maximum local de c_t , tel qu'illustré à la figure 1.

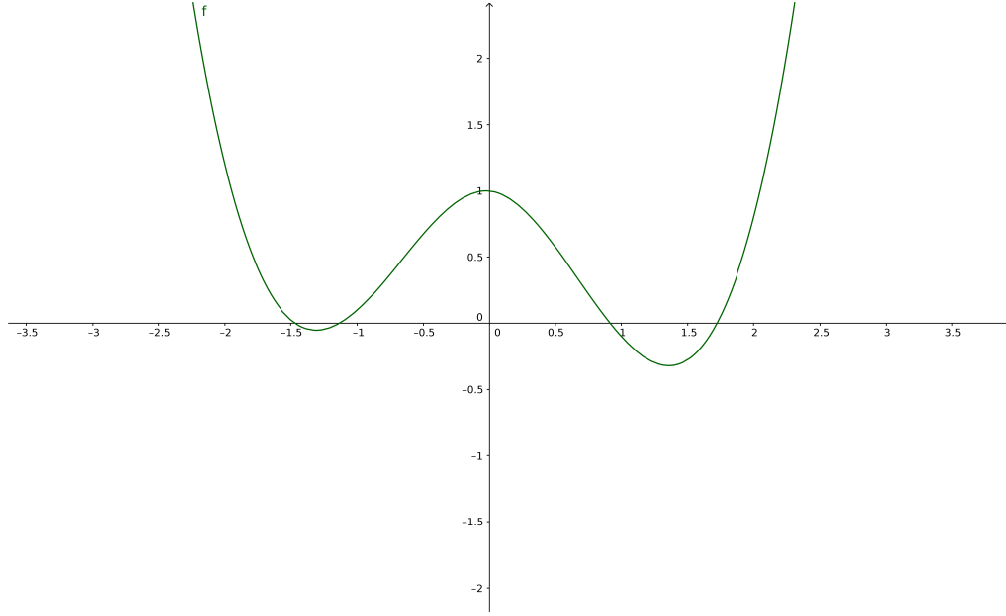


Figure 1 – Forme typique de c lorsque c'_t a trois racines

Pour s'assurer qu'on a des minima locaux, on utilise la dérivée seconde :

$$c''_t(d) = H + \frac{2}{\alpha} |d|.$$

On voit évidemment que dans les cas où $H > 0$ (les cas avec une seule racine), la racine est un minimum. En effet, $H > 0$, $\alpha > 0$ et $|d| > 0$ pour tout d . On a $c''_t(d) > 0$, donc on a un minimum local en vertu des conditions suffisantes d'optimalité (c'est-à-dire $c'_t(d) = 0$ et $c''_t(d) > 0$). Donc on a une procédure de sélection très simple lorsque $H > 0$. Si $g > 0$, alors on prend $d = \frac{-H + \sqrt{\delta_2}}{\frac{-2}{\alpha}}$ et sinon $d = \frac{-H + \sqrt{\delta_1}}{\frac{2}{\alpha}}$. Il nous reste à montrer que si $H < 0$, on a deux minima locaux.

- **Cas 1 : $H < 0$ et $g > 0$.** Les racines de c'_t sont respectivement

$$d_1 = \frac{-H - \sqrt{H^2 - 4\frac{g}{\alpha}}}{\frac{2}{\alpha}}, d_2 = \frac{-H + \sqrt{H^2 - 4\frac{g}{\alpha}}}{\frac{2}{\alpha}} \text{ et } d_3 = \frac{-H + \sqrt{H^2 + 4\frac{g}{\alpha}}}{\frac{-2}{\alpha}}.$$

$$c''_t(d_1) = H + \frac{2}{\alpha} \left| \frac{-H - \sqrt{H^2 - 4\frac{g}{\alpha}}}{\frac{2}{\alpha}} \right| = \underbrace{H}_{<0} + \underbrace{\left| -H - \sqrt{H^2 - 4\frac{g}{\alpha}} \right|}_{< |H|} < 0,$$

donc d_1 est un maximum local.

$$c_t''(d_2) = H + \frac{2}{\alpha} \left| \frac{-H + \sqrt{H^2 - 4\frac{g}{\alpha}}}{\frac{2}{\alpha}} \right| = \underbrace{H}_{<0} + \underbrace{\left| -H + \sqrt{H^2 - 4\frac{g}{\alpha}} \right|}_{> |H|} > 0,$$

donc d_2 est un minimum local.

$$c_t''(d_3) = H + \frac{2}{\alpha} \left| \frac{-H + \sqrt{H^2 + 4\frac{g}{\alpha}}}{\frac{-2}{\alpha}} \right| = \underbrace{H}_{<0} + \underbrace{\left| -H + \sqrt{H^2 + 4\frac{g}{\alpha}} \right|}_{> |H|} > 0,$$

donc d_3 est un minimum local.

Comme $d_1 - d_2 = \frac{-H - \sqrt{H^2 - 4\frac{g}{\alpha}}}{\frac{2}{\alpha}} - \frac{-H + \sqrt{H^2 - 4\frac{g}{\alpha}}}{\frac{2}{\alpha}} = \frac{-2\sqrt{H^2 - 4\frac{g}{\alpha}}}{\frac{2}{\alpha}} < 0$ et que $d_3 < 0$. Ainsi, on a bel et bien $d_3 < d_1 < d_2$. Le maximum local est encadré par les deux minimas locaux.

En conséquence, lorsque $H < 0$ et $g > 0$, la direction de descente est $\frac{-H + \sqrt{H^2 - 4\frac{g}{\alpha}}}{\frac{2}{\alpha}}$ ou $\frac{-H + \sqrt{H^2 + 4\frac{g}{\alpha}}}{\frac{-2}{\alpha}}$.

- **Cas 2 :** $H < 0$ et $g < 0$. Les racines de c_t' sont

$$d_1 = \frac{-H + \sqrt{H^2 - 4\frac{g}{\alpha}}}{\frac{2}{\alpha}}, d_2 = \frac{-H - \sqrt{H^2 + 4\frac{g}{\alpha}}}{\frac{-2}{\alpha}} \text{ et } d_3 = \frac{-H + \sqrt{H^2 + 4\frac{g}{\alpha}}}{\frac{-2}{\alpha}}.$$

$$c_t''(d_1) = H + \left| -H + \sqrt{H^2 - 4\frac{g}{\alpha}} \right| > 0,$$

donc d_1 est un minimum local.

$$c_t''(d_2) = H + \left| -H - \sqrt{H^2 + 4\frac{g}{\alpha}} \right| < 0,$$

donc d_2 est un maximum local.

$$c_t''(d_3) = H + \left| -H + \sqrt{H^2 + 4\frac{g}{\alpha}} \right| > 0,$$

donc d_3 est un minimum local. De plus, les racines s'ordonnent de la même manière que précédemment ($d_3 < d_2 < d_1$). Ainsi, si $H < 0$ et $g < 0$, la direction de descente est $\frac{-H + \sqrt{H^2 - 4\frac{g}{\alpha}}}{\frac{2}{\alpha}}$ ou $\frac{-H + \sqrt{H^2 + 4\frac{g}{\alpha}}}{\frac{-2}{\alpha}}$.

Le choix de la direction de descente se fait donc de la manière suivante : on pose $\delta_1 = H^2 - 4\frac{g}{\alpha}$ et $\delta_2 = H^2 + 4\frac{g}{\alpha}$. On observe qu'il est possible

que $H^2 - 4\frac{g}{\alpha}$ ou $H^2 + 4\frac{g}{\alpha}$ soit négatif. Toutefois, si c'est le cas, l'autre sera positif et nous donnera certainement un minimum global de c_t . On utilisera donc la direction issue de cette racine.

Nous avons donc une procédure de sélection lorsque $H < 0$ on choisit $d \in \left\{ \frac{-H+\sqrt{\delta_2}}{\frac{-2}{\alpha}}, \frac{-H+\sqrt{\delta_1}}{\frac{2}{\alpha}} \right\}$ tel que $c_t(d)$ a la plus petite valeur possible. Or, comme nous l'avons vu avec les méthodes de région de confiance, il est possible de choisir la meilleure direction de descente sans évaluer $c_t(d)$.

Le signe de g nous indique la racine qui minimise $c_t(d)$ globalement.

- **Cas 1 :** $g < 0$, illustré à figure 2.

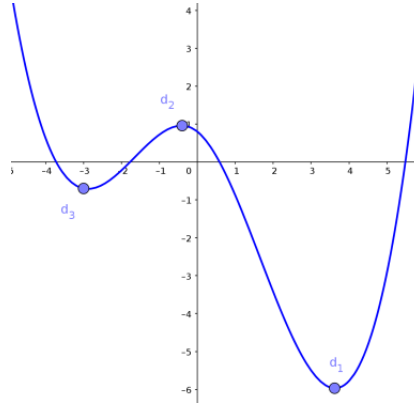


Figure 2 – Forme de c lorsque $H < 0$ et $g < 0$

Donc, si $g < 0$ et $H < 0$, alors le minimum global (donc la direction) est $\frac{-H+\sqrt{\delta_1}}{\frac{2}{\alpha}}$.

- **Cas 2 :** $g > 0$, illustré à la figure 3.

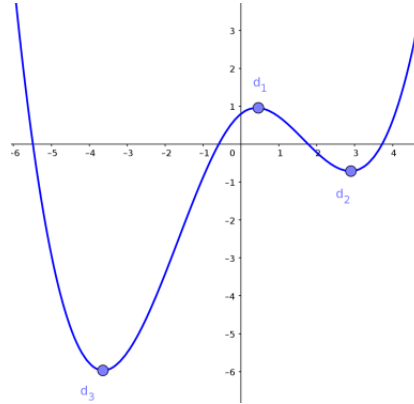


Figure 3 – Forme de c lorsque $H < 0$ et $g > 0$

Donc, si $g > 0$ et $H < 0$, alors le minimum global (donc la direction) est $\frac{-H+\sqrt{\delta_2}}{\frac{-2}{\alpha}}$.

□

3.3 Précision numérique

Un polynôme de la forme $ax^2 + bx + c$ admet ses racines en

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \Leftrightarrow x = \frac{-2c}{b \pm \sqrt{b^2 - 4ac}}.$$

Cette formule permet d'éviter les erreurs numériques lors de la programmation. En effet, on cherche à éviter l'addition de nombres de signes différents dans le code et on privilégie l'addition de nombres de mêmes signes. Par exemple, dans le choix de d , si $H > 0$ alors $-H + \sqrt{\delta_2}$ représente une addition de nombres de signes différents. En utilisant la formule alternative (celle à droite), on a $\frac{-2g}{H+\sqrt{\delta_2}}$ où l'on additionne seulement des nombres de mêmes signes.

Pour résumer, on énonce une procédure de sélection de direction qui traite tous les cas possibles (algorithme 5).

Algorithme 5 : Choix de la direction de descente pour ARC

```

1 si  $H > 0$  alors
2   si  $g > 0$  alors  $d = \frac{-2g}{H+\sqrt{\delta_2}}$ ;
3   sinon  $d = \frac{-2g}{H+\sqrt{\delta_1}}$ ;
4 sinon
5   si  $g < 0$  alors  $d = \frac{-H+\sqrt{\delta_1}}{\frac{-2}{\alpha}}$ ;
6   sinon  $d = \frac{-H+\sqrt{\delta_2}}{\frac{-2}{\alpha}}$ ;
7  $d_R \leftarrow d$ 
```

On remarque que l'analyse préalable rend le choix de la direction de descente très simple.

4 Tests avec le code

4.1 Algorithme de région de confiance

Les changements apportés à la procédure, c'est-à-dire trouver le minimiseur global de q_t sans utiliser d'évaluation de q_t) sont plutôt de nature esthétique, ils n'affectent pas les performances de l'algorithme, mais facilite l'écriture du code.

4.2 Algorithme de régularisation cubique adaptative

La procédure de choix de d a été incorporée dans un code existant en Julia [BKSE12] où le choix de d était plus arbitraire ("ancienne version" présentée

en annexe B) pour comparer les deux versions et ainsi valider l'implémentation de l'algorithme 5. Cet algorithme a été testé avec les problèmes d'optimisation unidimensionnelle de la bibliothèque *OptimizationProblems* voir [AMP], [A.S95] et [Dus15]. Notre outil de comparaison est le nombre d'évaluations des fonctions ϕ , ϕ' et ϕ'' lorsqu'on utilise les modèles de Newton.

Commençons par regarder ARC avec la variante de Newton ($H = \phi''(t)$). Les résultats sont présentés dans l'annexe A. Dans la table 4, on remarque que le nombre d'évaluations de fonctions avec la nouvelle version est toujours plus petit ou égal à celui de l'ancienne version. Les problèmes pour lesquels le nombre d'évaluations est strictement plus petit sont AMPGO20, Shpak2 et Shpak6.

Or, on remarque qu'avec le même algorithme, si on change le paramètre r_1 (qui gère si on réduit ou non l'importance de α), nos résultats vont changer. Dans l'exemple précédent, nous avons une valeur de $r_1 = 0.1$, mais si l'on prend une valeur de $r_1 = 0.4$, on obtient des résultats différents (voir table 4). Pour le problème Shpak2, l'ancienne version de l'algorithme a besoin de moins d'évaluations de fonctions. Pour les problèmes Shpak6 et AMPGO20, la nouvelle version a des meilleures performances.

On observe un phénomène similaire pour les algorithmes sécantes où le choix de r_1 influence beaucoup l'efficacité de la nouvelle version de l'algorithme. Par exemple, avec ARC-Sécante et $r_1=0.1$ (table 5), on observe que la nouvelle version est meilleure pour les problèmes Shpak6, Shpak2 et Shpak1. Or, l'ancienne version est meilleure pour les problèmes AMPGO02 et Dus2_9. Si on prend $r_1 = 0.4$, on obtient des résultats différents. Avec ce nouveau paramètre, l'ancienne version est plus efficace pour le problème AMPGO05. Toutefois la nouvelle version est meilleure pour les problèmes AMPGO02, AMPGO03, AMPGO08 et Shpak2.

Ainsi, le choix du paramètre r_1 est un facteur qui influence la performance des deux versions de l'algorithme. La nouvelle version de la procédure de sélection de d reste toutefois meilleure dans la majorité des cas. Il n'existe pas de choix optimal pour r_1 et r_2 . De manière générale on prend $r_1 = 0.25$ et $r_2 = 0.75$ ou $r_1 = 0.3$ et $r_2 = 0.7$.

5 Conclusion

Pour conclure, nous avons vu le fonctionnement des méthodes de région de confiance ainsi que des méthodes de régularisation cubique adaptative en dimension un. Ces méthodes approximent une fonction objectif ϕ par une fonction quadratique (q_t) ou cubique (c_t) et minimisent ces dernières. Nous avons montré qu'il est possible de trouver des minima globaux de ces fonctions de manière simple et efficace en utilisant seulement le signe des coefficients de degré un et deux de q_t et c_t . Évidemment, ces méthodes ne sont pas restreintes aux polynômes de degré deux ou trois, ainsi il est possible de prendre des polynômes de degré supérieur. Cependant, des outils de calcul de racines automatisés seront nécessaires pour déterminer les minima globaux. Le perfectionnement des

algorithmes de minimisation unidimensionnelle compte d'autres applications en optimisation. Entre autres, les méthodes de région de confiance et ARC ont été adaptées pour la recherche linéaire. Les méthodes ARC et TR sont également adaptables au cas multidimensionnel.

Références

- [AMP] 1-D Test Function. http://infinity77.net/global_optimization/test_functions_1d.html. Accédé le : 2017-05-08.
- [A.S95] A.SHPAK : Global optimization in one-dimensional case using analytically defined derivatives of objective function. *Computer Science Journal of Moldova*, 3(2):168–184, 1995.
- [BKSE12] Jeff BEZANZON, Stefan KARPINSKI, Viral SHAH et Alan EDELMAN : Julia : A Fast Dynamic Language for Technical Computing. *In Lang.NEXT*, avril 2012.
- [Dus15] J.-P. DUSSAULT : *Optimisation mathématiques avec applications en imagerie*. Sherbrooke, Québec, Canada, 2015.
- [XZ01] Chengxian XU et Jianzhong ZHANG : A Survey of Quasi-Newton Equations and Quasi-Newton Methods for Optimization. *Annals of Operations Research*, 103(1):213–234, 2001.
- [ZX01] Jianzhong ZHANG et Chengxian XU : Properties and numerical performance of quasi-Newton methods with modified quasi-Newton equations. *Journal of Computational and Applied Mathematics*, 137(2):269–278, 2001.

SAMUEL GOYETTE

DÉPARTEMENT DE MATHÉMATIQUES, UNIVERSITÉ DE SHERBROOKE

Courriel: Samuel.Goyette@USherbrooke.ca

A Résultats numériques pour ARC-Nwt et ARC-Sec

On présente les résultats numériques de ARC-Nwt et ARC-Sec. On a appliqué ces algorithmes à une collection de problèmes avec différentes valeurs de r_1 pour illustrer que ce paramètre n'a pas d'influence sur la convergence. Ainsi, l'algorithme modifié est plus performant que sa version standard.

Table 4 – Nombre d'évaluations de fonction avec ARC-Nwt

Nom du problème	$r_1 = 0.1$		$r_1 = 0.4$	
	Ancienne version	Nouvelle version	Ancienne version	Nouvelle version
AMPGO02	20	20	22	22
AMPGO03	24	24	24	24
AMPGO04	18	18	18	18
AMPGO05	24	24	24	24
AMPGO06	6	6	6	6
AMPGO07	20	22	22	22
AMPGO08	15	15	15	15
AMPGO09	15	15	15	15
AMPGO10	12	12	12	12
AMPGO12	6	6	6	6
AMPGO18	15	15	15	15
AMPGO20	31	28	31	28
AMPGO22	26	26	25	25
DUS2 1	114	114	114	114
DUS2 3	12	12	12	12
DUS2 9	24	24	24	24
DUSCUBE	20	20	20	20
SHPAK1	18	18	17	17
SHPAK2	20	17	19	22
SHPAK3	15	15	21	21
SHPAK5	17	17	17	17
SHPAK6	23	20	26	20

Table 5 – Nombre d'évaluations de fonction avec ARC-Sec

Nom du problème	$r_1 = 0.1$		$r_1 = 0.4$	
	Ancienne version	Nouvelle version	Ancienne version	Nouvelle version
AMPGO02	16	24	18	16
AMPGO03	34	34	34	32
AMPGO04	12	12	12	12
AMPGO05	34	34	32	34
AMPGO06	4	4	4	4
AMPGO07	26	26	16	16
AMPGO08	34	34	30	28
AMPGO09	14	14	14	14
AMPGO10	4	4	4	4
AMPGO12	4	4	4	4
AMPGO18	12	12	12	12
AMPGO20	44	44	44	44
AMPGO22	14	14	14	14
DUS2 1	102	102	102	102
DUS2 3	14	14	14	14
DUS2 9	66	70	24	24
DUSCUBE	16	16	18	18
SHPAK1	18	16	16	16
SHPAK2	18	16	20	18
SHPAK3	28	28	28	28
SHPAK5	36	36	38	38
SHPAK6	38	36	36	36

B Ancienne méthode choix de d pour ARC

Nous présentons ici l'ancienne procédure de sélection du minimiseur global de c_t . Étant donné que cette méthode manque de finition et ne résulte pas d'un raisonnement mathématique que l'on peut trouver, elle n'est pas présentée dans le cadre principal de l'article. Malgré le fait que cette procédure ressemble à celle que l'on a déjà, elle est incomplète et utilise des évaluations de c_t .

Algorithme 6 : Choix traditionnel de la direction de descente pour ARC

```

1   $\delta_0 = H^2 - 4\frac{g}{\alpha}$ ;
2  si  $\delta_0 < 0$  alors  $\delta_0 = H^2 + 4\frac{g}{\alpha}$ ;
3  si  $H < 0$  alors
4     $\left\lfloor d_{Rp} = \frac{-H + \sqrt{\delta_0}}{\frac{2}{\alpha}} \right.$ 
5  si  $H \geq 0$  alors
6     $\left\lfloor d_{Rp} = \frac{-2g}{H = \sqrt{\delta_0}} \right.$ 
7   $d_R n = \frac{-H + \sqrt{\delta_0}}{\frac{2}{\alpha}}$ ;
8  si  $c_t(d_R n) < c_t(d_{Rp})$  alors  $d_R = d_R n$ ;
9  sinon  $d_R = d_{Rp}$ ;
10 On retourne  $d_R$ 
```

La revue **CaMUS** souhaite remercier ces commanditaires, sans qui l'achèvement de ce projet aurait été beaucoup plus ardu. Pour ce volume, nous avons eu l'appui de trois généreux partenaires :

L'Association Générale des Étudiantes et Étudiants en Sciences



Le Regroupement des étudiants-chercheurs en sciences de l'Université de Sherbrooke



Le fonds institutionnel de soutien aux activités étudiantes de l'Université de Sherbrooke



CaMUS est une revue mathématique publiée par le Département de mathématiques de l'Université de Sherbrooke. Le but de ces cahiers est de permettre aux étudiants de présenter leurs travaux effectués dans le cadre d'activités comme les stages de recherche du premier cycle, les présentations au Club Mathématique et les cours d'initiation à la recherche. Les auteurs sont principalement des étudiants du premier cycle, notamment en mathématiques et au Baccalauréat en enseignement au secondaire avec profil en mathématiques. Cette revue est publiée à l'intention de tous ceux qui s'intéressent aux mathématiques, à l'Université de Sherbrooke et ailleurs. Ses objectifs sont de favoriser :

- l'apprentissage de la rédaction d'articles dès le premier cycle,
- l'amélioration de la communication écrite,
- le développement de la rigueur d'expression,
- la motivation des étudiants pour la recherche en général.

Le comité de rédaction de CaMUS est composé des étudiants :

- Étienne Bégin
- Antoine Bergeron
- Jean-Philippe Morissette (rédacteur gérant)
- David Racicot
- Fanny Rancourt
- Adam Salvail (rédacteur gérant)
- Hipolito Treffinger

et des professeurs :

- Taoufik Bouezmarni (rédacteur en chef)
- Tomasz Kaczynski

Information aux auteurs : Les articles doivent constituer des présentations originales mais on ne demande pas qu'ils contiennent des résultats originaux : pour la prépublication de ces derniers, la série de Rapports de recherche du Département de mathématiques est un meilleur médium. Les articles doivent être rédigés en \LaTeX et soumis suivant les instructions données à la page Web de CaMUS :

<http://camus.dmat.usherbrooke.ca>

Abonnement, achat d'exemplaires et autres renseignements : CaMUS est une publication électronique sans frais avec un nombre limité d'exemplaires imprimés à vendre. Pour des informations sur le prix courant, le mode de paiement ou pour nous contacter, rendez-vous à l'adresse Web indiquée ci-dessus.

CaMUS · Département de mathématiques · Faculté des sciences · Université de Sherbrooke · 2500, boulevard de l'Université · Sherbrooke (Québec), Canada J1K 2R1